

BAYESIAN APPROACH TO QUALITY CONTROL

by

PRAKASH VAMAN JOSHI

B.E. (Electrical)

University of Poona, India (1974)

A MASTER'S REPORT

submitted in partial fulfillment of the
requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

Kansas State University

Manhattan, Kansas

1976

Approved by:


Major Professor

LD
2668
R4
1976
J67
C.2
Document

ii

205

TABLE OF CONTENTS

CHAPTER		page
1	INTRODUCTION.....	1
1.1	Review.....	1
1.2	Defects.....	8
1.3	Bayesian theory.....	8
1.4	Bayes' theorem for the posterior density function of a parameter.....	10
1.5	Bayes' estimates.....	11
1.6	The Binomial case.....	14
1.7	Advantages and disadvantages of Bayesian analysis.....	15
2	ARTICLES.....	17
	<u>ARTICLE #1:</u> The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs.....	17
2.1.1	The compound hypergeometric distribution.....	21
2.1.2	The hypergeometric distribution.....	23
2.1.3	The binomial distribution.....	24
	<u>ARTICLE #2:</u> Bayesian single sampling attribute plans for continuous prior distributions.....	25
	<u>ARTICLE #3:</u> Serial sampling acceptance scheme derived from Bayes' theorem.....	29
2.3.1	Specification of model.....	29
	<u>ARTICLE #4:</u> Bayesian prediction and population size assumptions.....	35
2.4.1	Infinite population.....	35
2.4.2	Finite population.....	36
2.4.3	General case.....	38
	<u>ARTICLE #5:</u> A new prior distribution for attributes sampling.	39
	<u>ARTICLE #6:</u> An empirical Bayes' estimator for P (success) in the binomial distribution.....	46
2.6.1	Introduction.....	46
2.6.2	Bayes' estimator.....	47
2.6.3	Empirical Bayes' estimator.....	48
	<u>ARTICLE #7:</u> Bayesian solution of the single sample inspection scheme.....	52
2.7.1	Equation for the neutral line.....	52
2.7.2	Equation for optimal sample size.....	53
2.7.3	Illustrative example.....	56

**THIS BOOK
CONTAINS
NUMEROUS PAGES
WITH DIAGRAMS
THAT ARE CROOKED
COMPARED TO THE
REST OF THE
INFORMATION ON
THE PAGE.**

**THIS IS AS
RECEIVED FROM
CUSTOMER.**

CHAPTER		page
	<u>ARTICLE #8: Bayesian estimation for the proportions in a mixture of distributions.....</u>	58
	2.8.1 The likelihood function.....	58
	2.8.2 Bayesian estimation for p.....	60
3	SHORT SUMMARIES OF IMPORTANT ARTICLES.....	62
4	A LIST OF ARTICLES FOR ADDITIONAL READING.....	84
	4.1 Biometrika.....	84
	4.2 Technometrics.....	86
	4.3 Journal of American Statistical Association.....	86
	4.4 Annals of Mathematical Statistics.....	88
	4.5 Journal of Royal Statistical Society.....	89
	4.6 Miscellanea.....	89
	BIBLIOGRAPHY.....	91
	BIBLIOGRAPHY (Additional Reading).....	96

CHAPTER 1

INTRODUCTION

1.1 Review:

This report is primarily concerned with the Bayesian approach to various aspects of quality control. By way of introduction, we will briefly discuss both quality control and Bayesian theory.

First, acceptance sampling plans will be discussed. Inspection is done at various steps in manufacturing. It can be carried out for the raw material or at various stages while manufacturing, as well as for the final product. The final product may be inspected by the manufacturer himself or the customer may inspect the product at the time of purchase. This inspection is generally carried out on the sampling basis. Sampling inspection is done for various reasons. A product may be destroyed while testing or the cost of 100% inspection may be excessive. Hence all acceptance tests are done on sampling basis. Due to modern acceptance sampling methods, it is possible to carry out better quality improvement than that which might be possible by 100% inspection. If a large percentage of defective items are found in a batch, it is proper to reject the whole batch instead of rejecting individual items. By this method striking quality improvement can be carried out [Grant & Leavenworth].

The following symbols are generally used in relation with sampling acceptance plans -

N = number of items in a given lot.

n = number of items in a sample.

M = number of defective items in a given lot of size N .

m = number of defective items in a given sample of size n .

c = acceptance number, the maximum allowable number of defective items in a sample of size n .

p = sample fraction defective = m/n .

p' = true process average fraction defective = M/N .

\bar{p} = average fraction defective for several observed samples.

P_a = probability of acceptance of a lot.

β = consumer's risk, the probability of accepting a bad product.

α = producer's risk, the probability of rejecting a good product.

The acceptance sampling plan is called a single sampling plan when the decision is based on the basis of one sample only. For all single sampling plans these three numbers should be specified; the number of items N in the lot from which the sample is to be drawn, the number of items n in the random sample drawn from the lot and the acceptance number c .

In judging various acceptance sampling plans it is desirable to compare their performance over a range of possible quality levels of submitted product. A true picture of this can be given by operating characteristic (OC) curve. For any given fraction defective p' in a submitted lot, the OC curve shows the probability P_a that such a lot will be accepted by the given sampling plan. In many acceptance sampling procedures at least one defective item is permitted in the sample. It has been found that the operating characteristics of plans with acceptance number greater than zero are better than those having acceptance number equal to zero. Large acceptance numbers invariably involve large sample sizes when the accepting lots are stated to have a fixed percentage of defective items. One major advantage of having large sample size is that they have greater ability to distinguish between acceptable and unacceptable lots. Since a large sample size may

involve higher inspection costs, there should be some compromise between acceptance number and the sample size.

A double sampling plan involves the possibility of postponing the decision until a second sample is taken. In the single sampling plan the decision is based on one sample only. But in the double sampling plan, if the first sample is neither good nor bad, the decision is made on the basis of first and second samples combined. Double sampling plans generally involve less total inspection than the single sampling plan. They also have the advantage that a second chance is given to doubtful lots.

The additional symbols generally used in connection with double sampling are as follows:

n_1 = number of items in the first sample.

c_1 = acceptance number for first sample.

n_2 = number of items in the second sample.

$n_1 + n_2$ = combined number of items in the two samples.

c_2 = acceptance number for the two samples combined.

The term multiple sampling is used when three or more samples are inspected to decide whether to accept or reject the given lot. The decision on acceptance or rejection must be reached after a stated number of samples. The term sequential sampling is used when a decision can be made after each item has been inspected and when there is no restriction on the total number of items to be inspected.

The point $100 p_{0.10}$ is defined as the Lot Tolerance Per Cent Defective (LTPD), the percent defective that has the probability of acceptance of 10%.

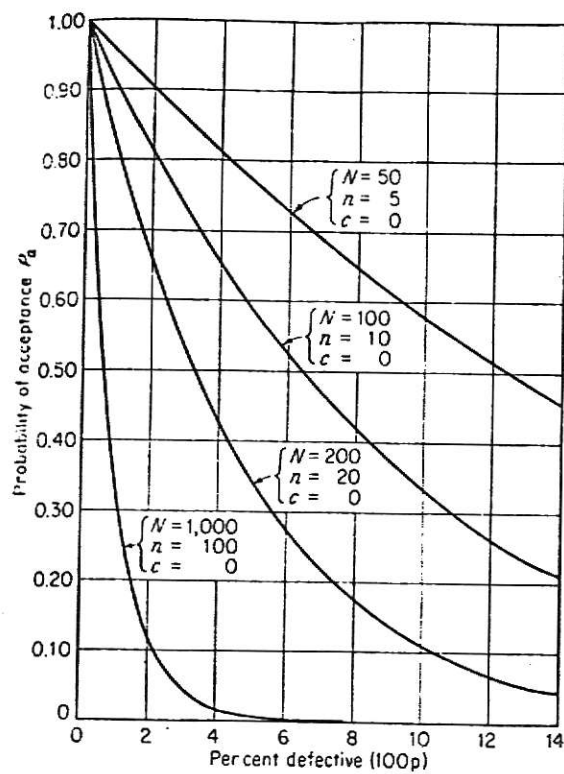


Fig. 1 Comparison of OC curves for four sampling plans involving 10% samples. [Grant & Leavenworth]

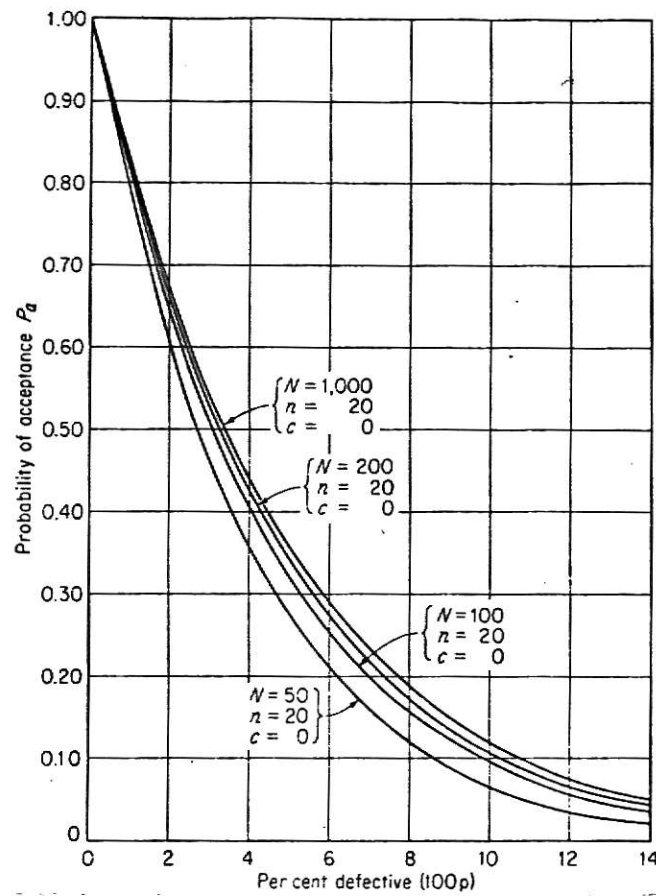


Fig. 2 Comparisons of OC curves for four sampling plans involving samples of 20, each with acceptance number 0. [Grant & Leavenworth]

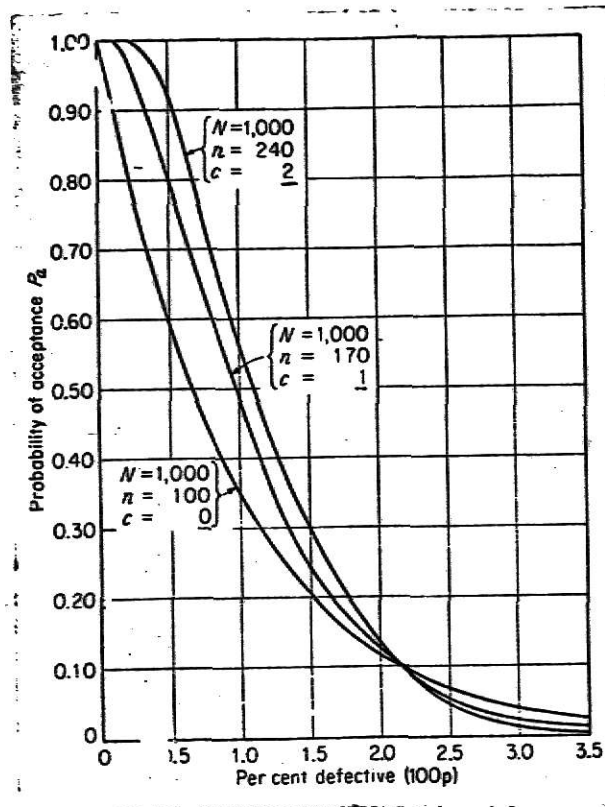


Fig. 3 OC curves for three sampling plans having a 0.10 probability of acceptance of a 2.2% defective lot. [Grant & Leavenworth]

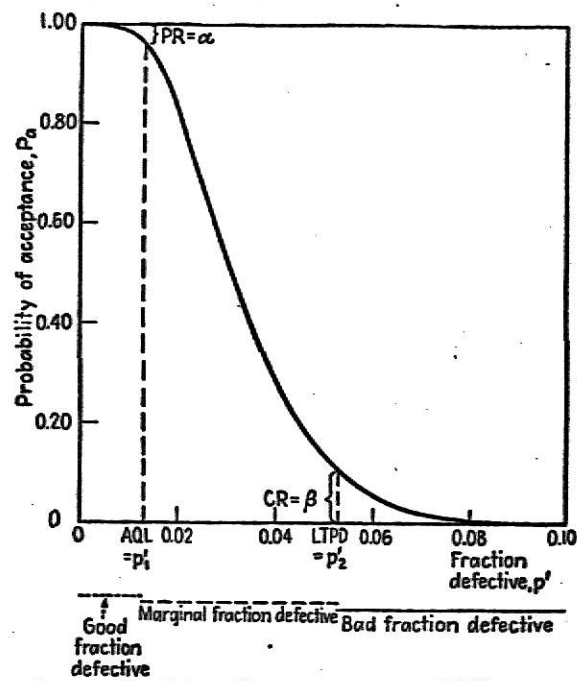


Fig. 4 An OC curve showing the important notations associated with it ($N = 3000$, $n = 150$, $c = 4$). [Burr]

All sampling plans in the Dodge-Romig tables have the main purpose of minimizing the Average Total inspection (ATI) by taking into account sampling inspection and screening inspection of rejected lots. The average outgoing quality limit, AOQL is defined as the maximum possible value of the average per cent defective in the outgoing product assuming that rejected lots are inspected 100%.

In many private industries, an AQL (acceptable quality level) system is widely used. The AQL is defined as the percent defective that is considered acceptable as a process average.

1.2 Defects:

Defects can be divided into 3 major groups; critical, major and minor (Grant and Leavenworth). They are defined as follows:

Critical defect: A critical defect is defined as a defect that judgement and experience indicate is likely to result in hazardous or unsafe conditions for individuals depending on the product.

Major defect: A major defect is a defect that is likely to result in failure, or to reduce the usability of the unit of product for its intended purpose.

Minor defect: A minor defect is a defect which is not likely to reduce the usability of the unit of product for its intended purpose nor it has any direct bearing on the effective use or operation of the unit.

1.3 Bayesian theory:

A large part of the work of an industrial engineer is concerned either with making decision under the conditions of uncertainty and partial ignorance that necessarily exist in many of his activities or with collecting and processing information that will be helpful in making such decisions.

The modern theory of Bayesian statistical inference and decisions deals with the development of techniques that are appropriate for making inferences and decisions in situations like this. The basis of the decision theory approach is Bayes' theorem which can be stated as follows:

Suppose that A_1, A_2, \dots, A_k are k mutually exclusive and exhaustive events, and suppose that B is any other event. Suppose that $P(A_i) > 0$ for $i = 1, 2, \dots, k$ and $P(B) > 0$. The following relation is known as Bayes' theorem:

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^k P(B|A_j)P(A_j)} \quad (1.3.1)$$

This relation provides a rule for computing the conditional probabilities $P(A_i|B)$ for $i = 1, 2, \dots, k$ from the conditional probabilities $P(B|A_i)$ and the probabilities $P(A_i)$ for $i = 1, \dots, k$. The values $P(A_i)$ for $i = 1, \dots, k$ are often called the "prior probabilities" of the events A_i , because they are the probabilities before it is known whether the event B has occurred. The values $P(A_i|B)$ for $i = 1, \dots, k$ are then called the 'posterior probabilities', because they are the relevant values after it is learned that B has occurred.

Bayes' theorem for a series of experiments can be developed as follows:

Let x be a random variable with conditional probability $f(x|\theta)$. Let θ be the parameter with discrete prior probability function $g(\theta)$. After a single observation X_1 , we can form the posterior probability of θ by the usual relationship: posterior = joint/marginal

$$\xi(\theta|x_1) = \frac{f(x_1|\theta)g(\theta)}{k(x)} = \frac{f(x_1|\theta)g(\theta)}{\sum_Z f(x_1|Z)g(Z)} \quad (1.3.2)$$

After a second observation x_2 (assumed independent of the first), we can form a new posterior, using as 'prior' the first stage posterior probability given by 1.3.2

$$\xi(\theta|x_1, x_2) = \frac{f(x_1, x_2|\theta)g(\theta)}{\sum_W f(x_1, x_2|W)g(W)} \quad (1.3.3)$$

The general formula for a series of observations x_1, x_2, \dots, x_n is

$$\xi(\theta|x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n|\theta)g(\theta)}{\sum_W f(x_1, x_2, \dots, x_n|W)g(W)} \quad (1.3.4)$$

1.4 Bayes' theorem for the posterior density function of a parameter:

Let the parameter W can take any value w in some interval $a < w < b$ and let the prior distribution of w be specified in terms of a density function $g(w)$ for $a < w < b$. We will assume that a discrete or a continuous observation X can be taken. For any given value w of W , $f(x|w)$ will denote either the probability function or the density function of X when $W = w$. In this case, the posterior density function $g(w|x)$ of W , after the value $X = x$ has been observed, is specified by the following equation:

$$g(w|x) = \frac{f(x|w)g(w)}{\int_a^b f(x|w')g(w')dw'} , \quad a < w < b \quad (1.4.1)$$

This is Bayes' theorem for densities.

1.5 Bayes' estimates:

Let x be considered the basic random variable under consideration. A sample is drawn yielding observations x_1, x_2, \dots, x_n . We will use these observations to construct a statistic y which will be useful for giving information about θ in the density function $f(x|\theta)$.

Let θ = the true value of the parameter.

$\hat{\theta} = w(y)$ = the estimator based on sample data.

If $\hat{\theta} \neq \theta$, a loss is incurred. It seems reasonable to assume that the loss function is a monotone increasing function of $\theta - \hat{\theta}$.

We will say that $L(\theta, \hat{\theta}) = L(\theta, w(y))$ is the loss incurred when θ is the true value of the parameter and we estimate it as $\hat{\theta} = w(y)$.

We define risk as average loss or expected loss.

$$\text{Risk} = R(\theta, w) = \int_{-\infty}^{\infty} L(\theta, w(y))h(y|\theta)dy \quad (1.5.1)$$

where $h(y|\theta)$ is the conditional probability.

Our main purpose here is to minimize the risk $R(\theta, w)$ expressed by equation 1.5.1. But we can not minimize this risk as it is expressed in terms of θ . Hence we have to express it in terms of average risk $\rho(g, w)$. This risk $R(\theta, w)$ depends explicitly on the true value of θ and implicitly on the structural form of $\hat{\theta} = w(y)$. Now we can find the average risk over all possible values of θ . Designating this average risk as $\rho(g, w)$, we have

$$\rho(g, w) = \int_{-\infty}^{\infty} R(\theta, w)g(\theta)d\theta \quad (1.5.2)$$

By expanding the right hand side, we get

$$\rho(g,w) = \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} L(\theta,w) \xi(\theta|y) d\theta \right] k(y) dy \quad (1.5.3)$$

where $\xi(\theta|y)$ is the posterior probability of θ .

Now we want to choose a $w(y)$ which minimizes $\rho(g,w)$ for all experimental outcomes y , so we should minimize the inner integral. We choose $w(y)$ to minimize

$$\begin{aligned} \theta(w) &= \int_{-\infty}^{\infty} L(\theta,w) \xi(\theta|y) d\theta \\ &= E[L(\theta,w(y)) | y] \end{aligned} \quad (1.5.4)$$

The estimator $w(y)$ so chosen is called the Bayes estimator of θ for the loss function $L(\theta,w)$. The corresponding average risk is Bayes risk.

Here is a list of commonly used loss functions and the corresponding Bayes' estimators. [Refer Table I on the next page]

In many cases, it is often required to choose a prior. Let the prior be chosen from a family with two characteristics [54].

- (1) The family should be analytically tractable in three respects:
 - (a) It should be reasonably easy to determine the posterior distribution resulting from a given prior distribution and a given sample,
 - (b) It should be possible to express in convenient form the expectations of some simple utility functions with respect to any member of family,
 - (c) The family should be closed in the sense that if a prior is a member of the family, the posterior will also be a member of the family.

Table I

$L(\theta, w)$	Bayes estimator of θ
$L = c(\theta - w)^2$	$E(\theta y)$
$L = m \theta - w $	w should divide the posterior distribution of θ into two equal parts
$L = a(\theta)(\theta - w)^2$	$w = \frac{E(a(\theta)\theta y)}{E(a(\theta) y)}$
$L = m_1(\theta - w), w < \theta$ $= -m_2(\theta - w), w > \theta$	w should divide the posterior distribution of θ into ratio $P_1 : P_2$ where $P_1 = \frac{m_2}{m_1 + m_2}$ and $P_2 = \frac{m_1}{m_1 + m_2}$
$L = C$ for $\theta \neq w$ $= 0$ for $\theta = w$	w should be the posterior mode.
$L = m(\theta - w), w < \theta$ $= 0, w > \theta$	$w = \infty$

- (2) The family should be rich, so that there will exist a member of the family capable of expressing the decision maker's prior information and beliefs. A prior distribution chosen according to these criteria is said to be conjugate to the conditional distribution.

1.6 The Binomial Case:

Let x be the random variable with probability function

$$f(x|n, \theta) = \binom{n}{x} \theta^x (1-\theta)^{n-x}, \quad x = 0, 1, 2, \dots, n. \quad (1.6.1)$$

Let θ be a realization of a random variable $\tilde{\theta}$ whose prior density can be represented by the beta density

$$g(\theta) = \frac{\theta^{a-1} (1-\theta)^{b-1}}{B(a, b)}, \quad 0 \leq \theta \leq 1 \quad (1.6.2)$$

The joint distribution of x and θ is given by

$$f(x|n, \theta)g(\theta) = \frac{\binom{n}{x} \theta^x (1-\theta)^{n-x} \theta^{a-1} (1-\theta)^{b-1}}{B(a, b)} \quad (1.6.3)$$

where $x = 0, 1, \dots, n$ and $0 \leq \theta \leq 1$

The marginal distribution of x is given by

$$\begin{aligned} k(x) &= \int_0^1 \binom{n}{x} \frac{1}{B(a, b)} \theta^{x+a-1} (1-\theta)^{n-x+b-1} d\theta \\ &= \frac{\binom{n}{x}}{B(a, b)} B(x+a, n-x+b), \quad x = 0, 1, \dots, n. \end{aligned} \quad (1.6.4)$$

This is called beta-binomial or hyperbinomial distribution.

The posterior distribution of θ is given by

$$\begin{aligned}\xi(\theta | x) &= \frac{\binom{n}{x} \frac{1}{B(a,b)} \theta^{a+x-1} (1-\theta)^{b+n-x-1}}{\binom{n}{x} \frac{1}{B(a,b)} B(a+x, b+n-x)}, \quad 0 \leq \theta \leq 1 \\ &= \frac{\theta^{a+x-1} (1-\theta)^{b+n-x-1}}{B(a+x, b+n-x)}\end{aligned}\quad (1.6.5)$$

Thus the Bayesian estimator of θ for squared error loss function is given by

$$\hat{\theta} = E(\theta | x) = \frac{a+x}{(a+x) + (b+n-x)} = \frac{a+x}{a+b+n} \quad (1.6.6)$$

This is in comparison with classical estimator

$$\hat{\theta} = \frac{x}{n}$$

1.7 Advantages and disadvantages of Bayesian Analysis:

Advantages:

- (1) It enables us to combine current experimental data with past information.
- (2) It sometime gives useful estimates of θ when the classical estimator is useless.
- (3) Bayesian estimators often have smaller variance than classical estimators.
- (4) In the Bayesian analysis, if the hypothesis is "certain", then no matter what the experimental outcome is, the posterior hypothesis

is also "certain"; and if the hypothesis is untenable, then no matter what the experimental outcome is, the posterior hypothesis is also untenable.

$$\text{i.e. } P(H) = 0 \text{ implies } P(H|X) = 0$$

$$\text{and } P(H) = 1 \text{ implies } P(H|X) = 1$$

Disadvantages:

- (1) Bayesian estimates are often biased.
- (2) The prior may be difficult to determine in Bayesian analysis.
- (3) Bayesian analysis is hard to explain to the layman.

A brief review of report

The main purpose of this report is to see how Bayesian analysis is applicable in Quality Control. In all, this report contains 8 articles dealing with various aspects of Quality Control such as optimum sample size, acceptance number, costs for accepted or rejected lots, subclasses in stratified sampling, population size assumption, single sample inspection schemes etc. To quote a few results, it has been found out that optimum sample size is directly proportional to the square root of the lot size; it is also shown that for the prediction problem, the population may be assumed to be infinite or finite.

The report also gives a short summary of articles published by various authors in various journals that are concerned with Bayesian analysis. A list of articles for additional reading has also being given. In the end a bibliography is given.

This report covers the following journals over a period of 10 to 12 years. (1) Biometrika, (2) Technometrics, (3) Journal of Royal Statistical Society, (4) Journal of American Statistical Association, (5) Annals of Mathematical Statistics, (6) Journal of Quality Technology and (7) other miscellaneous journals which were cited.

Initially it was decided to cover a wide range of literature on Bayesian analysis, but due to the advanced and complicated nature of the subject, together with the lack of sufficient background, work has been severely restricted. Hence more complicated articles are either summarized or listed for additional reading.

CHAPTER 2

ARTICLES

ARTICLE #1THE COMPOUND HYPERGEOMETRIC DISTRIBUTION AND A SYSTEM OF SINGLE SAMPLING
INSPECTION PLANS BASED ON PRIOR DISTRIBUTIONS AND COSTS BY A. HALD [32].

The lot size N , the sample size n , the acceptance number c are the three numbers required to define a single sampling plan. The general decision rule is to accept the lot if the number of defective items in the lot are $\leq c$, and if they are $> c$; reject the lot. The possibility of acceptance for a lot with fraction defective $p = X/N$ is

$$P_a(p) = \sum_{x=0}^c \binom{X}{x} \binom{N-X}{n-x} / \binom{N}{n} \quad (2.1.1)$$

where X denotes the number of defective items in the lot (denoted as M in the introduction).

We get the operating characteristic (OC) curve by plotting P_a against fraction defective of the inspection lot. The two most important points on the OC-curve are generally called the producer's risk (α) and consumer's risk (β). Usually it is difficult to fix α and β on the OC curve in a proper manner as we have to consider the prior distribution as well as cost. Here are some questions which should be considered before fixing risk points. What fraction defective a consumer can accept without any complaint? What proportion of fraction defective will be intolerable for him? What is the prior distribution of lots that has been previously supplied to the market? What prior distribution do the suppliers of the market have? Even though we do come with a plan by considering the above mentioned questions, the

plan should be considered a random one as there is no perfect way in which producer's risk and consumer's risk and risk points can be chosen.

Denoting producer's risk point by p_1 and consumer's risk point by p_2 , we can write, according to Peach and Littauer (53)

$$[(n-c)p_1]/[(c+1)q_1] = F_{\alpha}$$

$$\text{and } [(n-c)p_2]/[(c+1)q_2] = F_{1-\beta} \quad (2.1.2)$$

where $q = 1-p$, $p\{F < F_{\alpha}\} = \alpha$, and the degrees of freedom for F are $2(c+1)$ and $2(n-c)$. The solution of these equations for $\alpha = 0.05$ and $\beta = 0.10$ has been carried out by Grubbs (Ann. Math. Stat., 20, 1949) who tabulated p_1 and p_2 as functions of (n,c) for $c = 0(1)9$ and $n = 1(1)150$.

To get a simpler solution, the Poisson distribution can be used, which gives a fairly good approximation.

$$2np_1 = \psi_{\alpha}^2$$

and

$$2np_2 = \psi_{1-\beta}^2 \quad (2.1.3)$$

or in another form

$$R = p_2/p_1 = \psi_{1-\beta}^2 / \psi_{\alpha}^2$$

and

$$n = \psi_{\alpha}^2 / 2p_1 = \psi_{1-\beta}^2 / 2p_2 \quad (2.1.4)$$

where the degrees of freedom for ψ^2 are $2(c+1)$. Tables to facilitate the solution of R and n have been given by Peach and Littauer (53).

Generally it is very difficult to explain the four quantities (p_1, α) and (p_2, β) to the inspector. At the most we can specify for him p_1 and p_2

on the basis of prior knowledge and techno-econo considerations. Then we still have the problem of selecting the two risks. This problem can be solved by selecting $\alpha = 0.05$ and $\beta = 0.10$.

It can be said that if the prior is distributed binomially with parameter p , then the number of defective items in a sample and in the remaining part of the lots are also binomially distributed with parameter p . In this case the average number of defectives in accepted lots will be $(N-n)p$. Here we will assume that the inspection cost per item is k and the loss incurred by accepting a defective item is unity. Then the loss without carrying out any inspection becomes Np and the cost of total (100%) inspection becomes Nk . If we assume that the quality is acceptable, then we should have $p < k$. The cost for lots that are accepted after doing sampling inspection is $nk + (N-n)p$ on the average. When a lot is rejected, it is assumed that it is entirely inspected, hence the cost of inspection becomes Nk . The average total cost for sampling inspection then becomes the weighted average of the two above mentioned costs for accepted and rejected lots respectively, which can be written as

$$nk + (N-n)[pP_a + k(1-P_a)] \quad (2.1.5)$$

This expression is always greater than Np for $p < k$.

Hence we can safely conclude that sampling inspection is more expensive than accepting a lot without inspection even after taking into consideration the losses incurred by accepting defective items. This is expected since the process is supposed to be acceptable.

The fundamental concept of the Dodge-Romig (16) system is to select a plan that minimizes the average total inspection (ATI) cost and simultaneously gives the required consumer protection. Assuming the cost of

inspecting an item as unity, the cost for an average inspection can be given as

$$I = n + (N-n)(1-P_a) \quad (2.1.6)$$

Here we can see that I is a function of n and c , as well as of p . If we average over the binomial prior distribution, we will get the same result where p will be displaced by \bar{p} , the process average. The optimum values of (n, c) are then those values that minimize the cost function $I(\bar{p})$ under the condition that $P_a(p_t) = 0.10$ where p_t = Lot Tolerance Percent Defective (LTPD).

We can approach the problem of defective items from a different angle. We can replace the defective items by good items and consider the additional cost as a part of production costs. In this way we limit ourselves initially to taking into account the inspection costs. Suppose we measure the damage caused by defective items in accepted lots economically. The loss from the consumer's point of view can be divided into many categories, such as price paid per item, costs of handling and identifying the defective item, and the costs of assembling and dis-assembling. However, if the producer clears the defective product as a good product after inspection the consumer may ask for replacement. In this case the producer will have to face service and replacement costs in addition to loss of goodwill in the market.

Hence the loss may change according to particular situation but we will consider the loss incurred by accepting a defective item as an economic unit. Consider the case where the lot is accepted without inspection. For a lot having p fraction defective, the total cost will

be Np . We assume that costs of rejected lots are directly proportional to $(N-n)$ and denote the costs of rejected lots per item by k_r . Thus if we sort the rejected lots, k_r is the sorting costs per item divided by the costs of accepting a defective item. Let k_s be the cost of sampling inspection. Here we have $k_r \leq k_s$. [Fig. 5]

The costs linked with lots of quality $p = X/N$ in sampling inspection are composed of two expressions

(1) the costs for accepted lots that are

$$nk_s + (X-x)(1) \text{ for } 0 \leq x \leq c,$$

and

(2) the costs for rejected lots that are

$$nk_s + (N-n)k_r \text{ for } c+1 \leq x \leq n, \quad (2.1.7)$$

where X is the number of defective items in the lot and x is the number of defective items in the sample. Now the probability of getting x defective items in a sample from a lot containing X defective items is

$$p\{x|X\} = \binom{X}{x} \binom{N-X}{n-x} / \binom{N}{n} \quad (2.1.8)$$

Hence the average cost for lots of quality $p = X/N$ becomes

$$K(n, c, p) = nk_s + \sum_{x=0}^c (X-x)p\{x|X\} + (N-n)k_r \sum_{x=c+1}^n p\{x|X\} \quad (2.1.9)$$

2.1.1 The compound hypergeometric distribution

From the conditional hypergeometric distribution $p\{x|X\}$ and the prior distribution $F_N(X)$, i.e. the probability that a lot of N items contains X defective items, we have the marginal probability

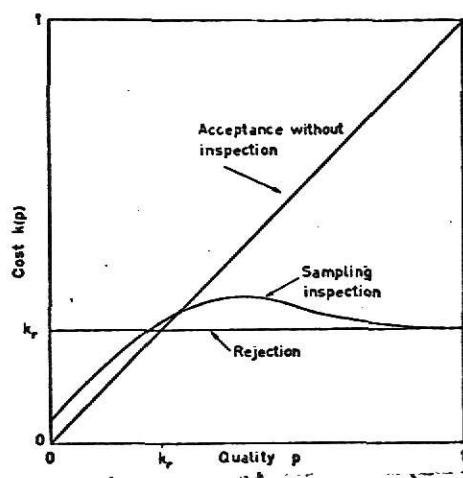


Fig. 5 Cost functions per item submitted. [32]

$$p\{X, x\} = F_N(X)p\{x/X\} \quad (2.1.10)$$

Let y denote the number of defective items in the part of lot which is not inspected. Then

$$y = X - x.$$

Here we have the probability

$$p(x, y) = F_N(x+y) \binom{n}{x} \binom{N-n}{y} / \binom{N}{x+y} \quad (2.1.11)$$

Hence the marginal distribution of x can be obtained as

$$g_n(x) = \binom{n}{x} \sum_{y=0}^{N-n} F_N(x+y) \binom{N-n}{y} / \binom{N}{x+y} \quad (2.1.12)$$

This distribution is called the compound hypergeometric distribution.

2.1.2 Hypergeometric distribution

Suppose we have a stock of M items containing $A = M\bar{p}$ defective items and from this stock, lots of N items are selected arbitrarily, then the prior distribution of X is given by

$$F_N(X; \bar{p}, M) = \frac{\binom{A}{X} \binom{M-A}{N-X}}{\binom{M}{N}} = \frac{\binom{N}{X} \binom{M-N}{A-X}}{\binom{M}{A}} \quad (2.1.13)$$

with mean $N\bar{p}$ and variance

$$\sigma_X^2 = N\bar{p}\bar{q} \frac{M-N}{M-1}$$

If we substitute $F_N(X; \bar{p}, M)$ into $p\{x, y\}$ derived in the compound hypergeometric distribution, we get

$$p\{x, y\} = F_n(x; \bar{p}, M) F_{N-n}(y; \bar{p}_x, M-n)$$

with

$$\bar{p}_x = \frac{A-X}{M-n} = (\bar{p} - \frac{x}{M}) / (1 - \frac{n}{M}) \quad (2.1.14)$$

This indicates that (a) the probability of x indicated by $g_n(x)$ is a hypergeometric distribution with the same parameters as prior distribution and (b) the probability of y for given x indicated by $p\{y|x\}$ is also a hypergeometric distribution with parameters \bar{p}_x and $M-n$.

2.1.3 The binomial distribution

Let the lots be produced with a process average equal to \bar{p} so that

$$F_N(X;\bar{p}) = \binom{N}{X} \bar{p}^X \bar{q}^{N-X} \quad (2.1.15)$$

with mean $N\bar{p}$ and variance $N\bar{p}\bar{q}$. Substituting this expression into $p\{x,y\}$ as in the previous case, we get

$$p\{x,y\} = F_n(x;\bar{p}) F_{N-n}(y;\bar{p}) \quad (2.1.16)$$

which indicates that x and y are independent and that both variables are binomially distributed with the same parameter as in the prior distribution.

ARTICLE #2

BAYESIAN SINGLE SAMPLING ATTRIBUTE PLANS FOR CONTINUOUS PRIOR DISTRIBUTIONS BY A. HALD [34].

In this article we will derive and discuss the fundamental properties of a system of single sampling attribute plans which are obtained by minimizing the average costs. The main assumptions made here are that sampling and decision costs are linear in lot size, sample size, and the number of defectives in the lot and the sample, that sampling is without replacement, and that the distribution of lot quality is a mixed binomial. This means that each lot is produced by a process in binomial control but the process average varies from lot to lot according to a frequency distribution.

Initially we will discuss a system of single sampling attribute plans by first minimizing the regret function

$$R(N,n,c) = n(p_s - p_m) + (N-n) \cdot \left\{ \int_0^{p_r} (p_r - p)Q(p)w(p)dp + \int_{p_r}^1 (p - p_r)P(p)w(p)dp \right\} \quad (2.2.1)$$

where N, n and c have the usual meanings; also $0 < p_s < 1$, $0 < p_r < 1$, $p_s \geq p_r > p_m$, $P + Q = 1$ and $P = P_a = P(p)$. Now P_a is given by

$$P_a = P(p) = \sum_{x=0}^c b(x,n,p) = B(c,n,p)$$

where $b(x,n,p) = \binom{n}{x} p^x q^{n-x}$

The two terms in the expression of R indicate sampling costs and average decision losses, respectively. This regret function is the

standardized average cost for a single sampling plan. Here p_s is sampling cost per item divided by the costs of accepting a defective item, p_r is the rejection cost per item divided by the costs of accepting a defective item and p_m is average unavoidable minimum cost. A detailed derivation of the model has been given by A. Hald [32 & 33].

Now the average unavoidable minimum cost p_m can be given by

$$p_m = \int_0^{p_r} p w(p) dp + \int_{p_r}^1 p_r w(p) dp = p_r - \int_0^{p_r} (p_r - p) w(p) dp \quad (2.2.2)$$

From (2.2.1) and (2.2.2) we get

$$R(N, n, c) = n(p_s - p_m) + (N - n)(p_r - p_m) - (N - n)\gamma(n, c) \quad (2.2.3)$$

$$\text{where } \gamma(n, c) = \int_0^1 B(c, n, p)(p_r - p)w(p) dp$$

Finally according to Hald it reduces to

$$c = np_r - \frac{1}{2} + p_r - q_r - p_r q_r w'(p_r) / w(p_r) \quad (2.2.4)$$

When the prior distribution is unimodal, it can be proved that $w'(p_r) < 0$, but the requirement here is that p_r should be larger than the mode, which is usually the case in practice.

For example consider $w(p) = p^{s-1} q^{t-1} / B(s, t)$ where $s > 0$, $t > 0$. It can be said that

$$c = np_r - \frac{1}{2} + (s+t)(p_r - \bar{p}) \quad (2.2.5)$$

where $E(p) = \bar{p} = s/(s+t)$. Eliminating $(s+t)$ and introduction the variance $\sigma_p^2 = \bar{p}q / (s+t+1)$ to equation (2.2.5), we get

$$c = np_r - \frac{1}{2} + (p_r - \bar{p})(\bar{p}q - \sigma_p^2) / \sigma_p^2 \quad (2.2.6)$$

From the prior information and experience it is relatively easy to estimate \bar{p} and σ_p while developing a system of sampling inspection plans. From equation (2.2.6) it can be said that the optimum acceptance number c is approximately a linear function of the sample size n .

According to Hald, we have

$$p_r - p_m - \gamma(n, nh) = -\alpha_1 n^{-1} - \alpha_2 n^{-2} + O(n^{-3}) \quad (2.2.7)$$

where $c = nh$, $\alpha_1 = -v/2$ and $24\alpha_2 = w + 2v - 10r'v' - 9rv'' + 12(v')^2 / w$

From (2.2.3) and (2.2.7), Hald obtained result that

$$R = n(p_s - p_m) - (N-n)(\alpha_1 n^{-1} + \alpha_2 n^{-2} + O(n^{-3})) \quad (2.2.8)$$

The equation $\frac{\partial R}{\partial n} = 0$ leads to

$$(N-n)\{-\alpha_1 - 2\alpha_2 n^{-1} + O(n^{-2})\} = n^2\{p_s - p_m + \alpha_1 n^{-1} + \alpha_2 n^{-2} + O(n^{-3})\} \quad (2.2.9)$$

$$\text{Solving for } n, \text{ Hald obtained } n = \lambda_1 \sqrt{N} + \lambda_2 + O(N^{-1/2}) \quad (2.2.10)$$

where $\lambda_1^2 = -\alpha_1 / (p_s - p_m)$ and $\lambda_2 = \alpha_2 / \alpha_1$, which leads to

$$\lambda_1^2 = rw / 2(p_s - p_m) \text{ and}$$

$$\lambda_2 = \{-3(1+4r)w + 4rr'w' + 9r^2w'^2 - 12r^2(w')^2/w\} / 12rw$$

To find the minimum average decision loss Hald used (2.2.9) to eliminate $(N-n)$ in the last term of (2.2.8) which then becomes

$(n - \lambda_1^2 - \lambda_2)(p_s - p_m) + O(n^{-1})$. Adding this to the first term, Hald got the minimum regret as

$$R_0(N) = (2n - \lambda_1^2 - \lambda_2)(p_s - p_m) + O(n^{-1}) \quad (2.2.11)$$

From the above equations it can be observed that the optimum sample size is approximately a linear function of the square root of the lot size N .

Hence as a summary we can say that the optimum sample size is approximately a linear function of the square root of the lot size and the optimum acceptance number is approximately a linear function of the sample size.

ARTICLE #3

SERIAL SAMPLING ACCEPTANCE SCHEME DERIVED FROM BAYES' THEOREM BY D. R. COX [14].

It is safe to assume in many industrial situations that if lots of items are produced sequentially, there will be some sort of correlation between those lots. Hence if we know the number of defectives in a random sample of fixed size taken from a certain batch to be screened and the number of defectives found in batches before and after, we can safely predict whether the batch is good or bad. Hence while making a decision on a particular batch, we have to take into consideration the sample outcome of that batch as well as the outcomes of the batches close to it. A stochastic process can be set up to represent serial sampling acceptance schemes and Bayes' theorem can be applied to it to obtain a screening rule.

2.3.1 Specification of model

We will consider the sequence of batches $(\dots, B_{n-1}, B_n, B_{n+1}, \dots)$ to be screened and let there be a corresponding sequence $(\dots, x_{n-1}, x_n, x_{n+1}, \dots)$ of number of defectives by counting the number of defectives in random samples of predetermined fixed size. We will assume that x_n has a Poisson distribution of mean m_n , where m_n is the true batch quality. We will also specify the stochastic process $\{m_n\}$. Bernard [5] has specified the following model

We will assume that a and b ($a < b$) are the only two positive values m_n has. We will call batches with $m_n = a$ as good batches and those with

$m_n = b$ as bad batches. Naturally we will accept good batches and reject bad batches. We will assume that the sequence $\{m_n\}$ has a simple Markov chain with transition matrix.

$$\begin{pmatrix} 1-t_a & t_a \\ t_b & 1-t_b \end{pmatrix}$$

This means that good and bad batches are independently geometrically distributed with mean run lengths $1/t_a$ and $1/t_b$. The prior odds that the batch is bad is then t_a/t_b . The sequence of true batch qualities is arbitrary if and only if $t_a + t_b = 1$.

Now we will enumerate the prior probability of a particular sequence of m 's. For the batches B_{n-1} and B_n , we have

$$\begin{aligned} \text{prob. } (m_{n-1} = m_n = a) &= \frac{t_b(1-t_a)}{t_a+t_b} \\ \text{prob. } (m_{n-1} = m_n = b) &= \frac{t_a(1-t_b)}{t_a+t_b} \end{aligned} \tag{2.3.1}$$

$$\text{prob. } (m_{n-1} = a, m_n = b) = \text{prob. } (m_{n-1} = b, m_n = a) = \frac{t_a t_b}{t_a+t_b}$$

For example, the probability that both batches are good is the probability that B_{n-1} is good viz. $\frac{t_b}{t_a+t_b}$, times the conditional probability that B_n is good given that B_{n-1} is good which is $1-t_a$. This model is reversible in time.

If we intend to screen the batch B_n on the basis of x_{n-1} and x_n , equation (2.3.1) gives prior probabilities of the various possible values

for m_{n-1} and m_n . We can write the probabilities according to Poisson distribution, e.g.

$$\text{prob. } (x_{n-1}, x_n \mid m_{n-1} = a, m_n = b) = e^{-a-b} \frac{a^{x_{n-1}} b^{x_n}}{x_{n-1}! x_n!} \quad (2.3.2)$$

The posterior probabilities of the various possible values for m_{n-1} and m_n can be derived from Bayes' theorem. The probability that B_n is bad is given by

$$\text{prob. } (m_n = b \mid x_{n-1}, x_n) = \frac{t_a(1-t_b)e^{-2b} a^{x_{n-1}} b^{x_n} + t_a t_b e^{-a-b} a^{x_{n-1}} b^{x_n}}{t_b(1-t_a)e^{-2a} a^{x_{n-1}} b^{x_n} + t_a t_b e^{-a-b} a^{x_{n-1}} b^{x_n}} \quad (2.3.3)$$

Now if the loss from rejecting a good batch is w_a and if the loss from accepting a bad batch is w_b , we will reject the batch if

$$\text{prob. } (m_n = b) > \frac{w_b}{w_a},$$

and accept if

$$\text{prob. } (m_n = b) < \frac{w_b}{w_a} \quad (2.4.4)$$

The choice is left open about what to do if $\text{prob. } (m_n = b) = \frac{w_b}{w_a}$. This is the optimum screening rule.

The rejection region can be plotted with the fact that if (x_{n-1}, x_n) is a rejection point, so also is any other point (x'_{n-1}, x'_n) for which $x'_{n-1} \geq x_{n-1}$ and $x'_n \geq x_n$. If a wrong decision is taken, the expected loss per batch is a function of a two point prior distribution (Fig. 6),

$$w_b P(\text{acceptance} \mid m_n = b) + w_a P(\text{rejection} \mid m_n = a).$$

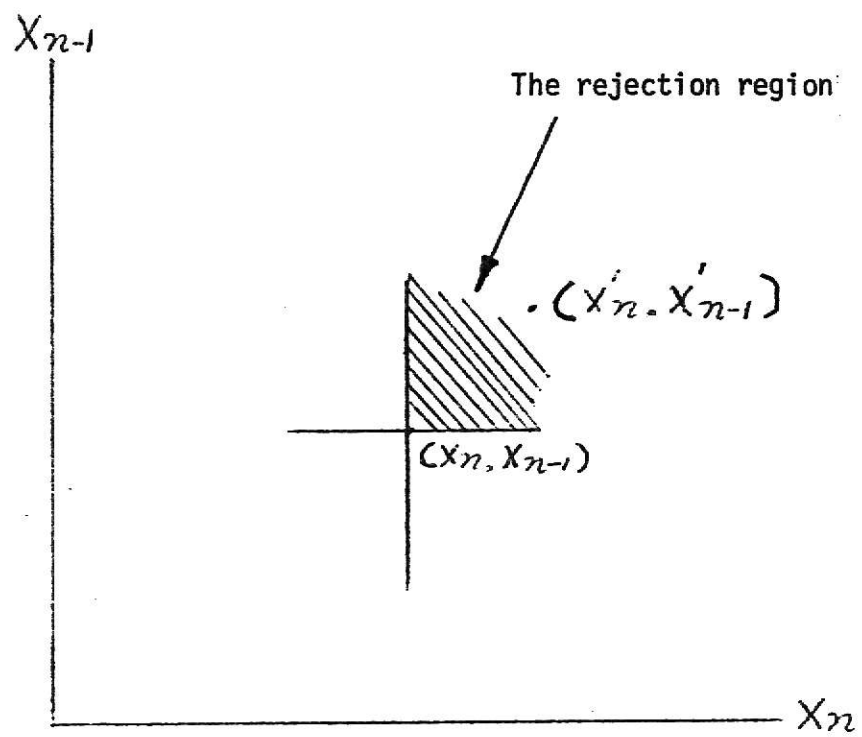


Fig. 6 The rejection region.

We will state some general properties of the scheme. We will assume a critical number r such that any batch having r or more defectives will be rejected irrespective of the outcome of previous batches. Note that as the process $\{m_n\}$ is Markovian, we will reject a batch B_n for all $(x_{n-1}, x_{n-2}, \dots)$ if and only if we would reject it if we knew that B_{n-1} is good i.e. $m_{n-1} = a$. Now, for $m_{n-1} = a$, the probability that B_n is bad is

$$\frac{(1-t_b)e^{-b}x_n}{t_b e^{-a}x_n}$$

$$\text{Thus } r \geq \log \left(\frac{t_b w_a e^b}{(1-t_b) w_a e^a} \right) / \log \left(\frac{b}{a} \right)$$

We may reject a batch B_n even when $x_n = 0$. In order to avoid this, we accept a batch when $x_n = 0$ even when it's clear that B_{n-1} is bad, i.e. that $m_{n-1} = b$. Hence we should have

$$\frac{(1-t_b)e^{-b}}{t_b e^{-a}} < \frac{w_b}{w_a} .$$

If there is an occurrence of consecutive bad batches, in order to improve quality of the batches, a change in the process can be made after the production of B_{n-1} batches. While screening B_n , it will be wrong to use x 's before x_n . The right procedure then is to calculate the posterior probability that B_n is bad by utilizing only x_n and subsequent x 's that can be made available while screening occurs.

In this discussion it has been assumed that the sample size is pre-determined and fixed. If the previous assumptions about the true batch quality remain valid, the calculation of the screening rule is rather easy if the sample size is varied.

ARTICLE #4

BAYESIAN PREDICTION AND POPULATION SIZE ASSUMPTIONS BY T. L. BRATCHER,
W. R. SCHCANY AND H. H. HUNT [10].

Bayesian methods aid in obtaining the distribution of the number of successes in a sample when the outcome of a previous sample is given. The main purpose of this article is to prove that the solution to the prediction problem doesn't depend on the population size. We can assume the population to be finite or infinite. We can also assume a uniform prior distribution on the proportion of successes and then Bayes' rule can be applied. It can be proved that infinite and finite cases have identical results.

2.4.1 Infinite population

Suppose we take a random sample of size m from an infinite population where θ is the unknown proportion of successes. Then Y , the number of successes in the sample has the binomial distribution

$$b_{\theta}(y,m) = \binom{m}{y} \theta^y (1-\theta)^{m-y}, \quad y = 0, 1, \dots, m \quad (2.4.1)$$

The main problem is to see how the number of successes X are distributed from a second sample of size n , given (y,m) . Let θ have a uniform prior distribution. Now, the posterior distribution of θ can be obtained by applying Bayes' rule,

$$f(\theta|y) = \frac{(m+1)!}{y!(m-y)!} \theta^y (1-\theta)^{m-y}, \quad 0 < \theta < 1 \quad (2.4.2)$$

The distribution of $f(X|\theta)$ is again binomial. By averaging out θ we have,

$$f(x|y) = \int_0^1 b_{\theta}(x,n) f(\theta|y) d\theta \quad (2.4.3)$$

because $f(x,\theta|y) = f(x|\theta) \cdot f(\theta|y)$

Now according to Raiffa and Schlaifer (3), the hypergeometric is given by

$$f(x|y) = \frac{\binom{x+y}{x} \binom{n-x+m-y}{n-x}}{\binom{n+m+1}{n}}, \quad x = 0, 1, \dots, n \quad (2.4.4)$$

2.4.2 Finite population

Now let the population be finite with size N and an unknown number of successes J . Here we will take the discrete uniform on the proportion J/N to get the prior distribution

$$f(J) = \frac{1}{N+1}, \quad J = 0, 1, \dots, N. \quad (2.4.5)$$

If we take a sample of size m , then the resulting probability function will be the hypergeometric

$$h_J(y,m) = \frac{\binom{J}{y} \binom{N-J}{m-y}}{\binom{N}{m}}, \quad y = 0, 1, \dots, m. \quad (2.4.6)$$

The joint probability will be

$$f(y,J) = \frac{\binom{J}{y} \binom{N-J}{m-y}}{\binom{N}{m} (N+1)}, \quad \begin{array}{l} y = 0, 1, \dots, m \\ J = y, y+1, \dots, N. \end{array} \quad (2.4.7)$$

Let s denote the remaining number of successes.

Hence

$$f(y,s) = \frac{\binom{s+y}{y} \binom{N-s-y}{m-y}}{\binom{N+1}{m} \binom{N-m+1}{m-y}}, \quad y = 0,1, \dots, m$$

$$s = 0,1, \dots, N-m. \quad (2.4.8)$$

$$\text{But } f(s) = \frac{1}{\binom{N-m+1}{m}}, \quad s = 0,1, \dots, N-m.$$

Now we can write

$$f(s|y) = \frac{\binom{s+y}{y} \binom{N-s-y}{m-y}}{\binom{N+1}{m+1}}, \quad s = 0,1, \dots, N-m. \quad (2.4.9)$$

Y has the uniform marginal probability

$$f(y) = \frac{1}{m+1}, \quad y = 0,1, \dots, m.$$

Hence if n is the sample size and x is the number of successes, then

$$h_s(x,n) = \frac{\binom{S}{x} \binom{N-m-S}{n-x}}{\binom{N-m}{n}}, \quad x = 0,1, \dots, n \quad (2.4.10)$$

Therefore, averaging out S, we have

$$f(x|y) = \sum_S h_s(x,n) f(s|y)$$

$$= \sum_{s=x}^{N-m-(n-x)} \frac{\binom{S}{x} \binom{N-m-S}{n-x} \binom{s+y}{y} \binom{N-s-y}{m-y}}{\binom{N-m}{n} \binom{N+1}{m+1}} \quad (2.4.11)$$

It can be shown that this expression is regardless of the assumed size of the population, N. According to Gould [27], we can write

$$\begin{aligned}
 f(x|y) &= \frac{\binom{x+y}{x} \binom{n+m-x-y}{n-x}}{\binom{n+m+1}{n} \binom{N+1}{n+m+1}} \sum_{s=x}^{N-m-n+x} \binom{S+y}{x+y} \binom{N-y-S}{n-x+m-y} \\
 &= \frac{\binom{x+y}{x} \binom{n+m-x-y}{n-x}}{\binom{n+m+1}{n}}, \quad x = 0, 1, \dots, n
 \end{aligned} \tag{2.4.12}$$

which is exactly same with the result obtained in 2.4.1.

2.4.3 General Case

The problem has been generalized to accommodate a broader class of prior distributions rather than the uniform distribution only. In the infinite population case, if we use a beta prior distribution with parameters v_1 and v_2 on θ and in the form given by Steck & Zimmer [63], if we use the negative hypergeometric prior distribution on J with parameters $a = v_2$ and $b = v_1$; it can be proved in a similar manner that the probability $f(x|y)$ is independent of the population size.

$$f(x|y) = \frac{\binom{v_1+x+y-1}{v_1+y-1} \binom{v_2+m+n-x-y-1}{v_2+m-y-1}}{\binom{m+n+v_1+v_2-1}{n}} \tag{2.4.13}$$

where $x = 0, 1, \dots, n$. By putting $v_1 = v_2 = 1$, the above expression reduces to results obtained in 2.4.1 and 2.4.2.

ARTICLE #5

A NEW PRIOR DISTRIBUTION FOR ATTRIBUTES SAMPLING BY W. K. CHIU [12].

In the Bayesian approach to attributes sampling, the beta distribution is the most widely used continuous prior distribution. The main reasons for using the beta distribution are (a) it is mathematically simple, and (b) the nonavailability of other suitable continuous prior distributions. However, the beta distribution also has some disadvantages (a) the beta density can't be expected to fit other prior densities, (b) the beta distribution doesn't have any practical model for which it can be used, (c) it is often necessary to have more than one type of prior distribution. Hence, in order to overcome the above mentioned difficulties, this article suggests a simple alternative prior distribution.

Let a process manufacturing batches of items have a normally distributed quality characteristics x with a mean μ and a variance of unity. The mean μ is presumed to have a normal prior distribution $N(m, \sigma^2)$ where m and σ^2 are known and do not depend on batch size. Let $\phi(x)$ be the density function and $\psi(x)$ be the distribution function of a standard normal distribution.

Let the batch have a fraction defective $p = \psi(x_0 - \mu)$ where μ is the mean quality and x_0 is that value below which an item is classified defective.

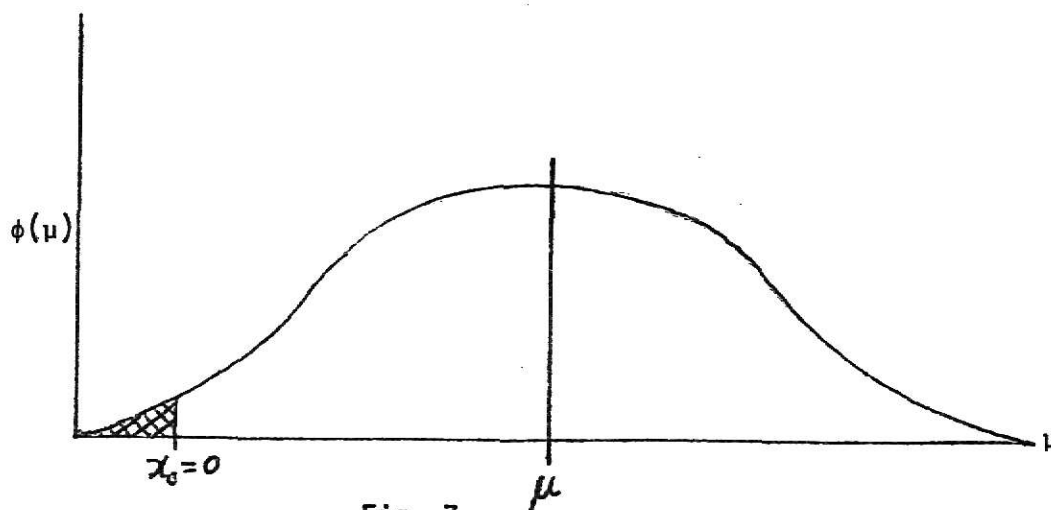


Fig. 7

Since $x \sim N(\mu, 1)$ and $\mu \sim N(m, \sigma^2)$

We have

$$p = \psi\left(\frac{x_0 - \mu}{1}\right), \quad \frac{dp}{d\mu} = -\phi(x_0 - \mu)$$

According to the author, x_0 is assumed to be equal to zero. The shaded area on the left side of x_0 represents the defective items. [Fig. 7] It is clear that the extent of this shaded area depends upon the value of μ . We can get the following density of the prior distribution of p , if we make a change of variables from μ to p ,

$$w(p) = \sigma^{-1} \exp \{ \mu^2/2 - (\mu - m)^2/(2\sigma^2) \} \quad (2.5.1)$$

where $p = \psi(x_0 - \mu)$, $0 \leq p \leq 1$.

We will call this prior distribution the "normal generated distribution".

The shape of the normal generated distribution for various values of σ^2 are shown in Fig. 8.

We can compare the shapes of a normal generated density and of a beta density when their means and variances are the same. The beta distribution with parameters a and b has the following density function:

$$f(p; a, b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a, b)} \quad (2.5.2)$$

where $a > 0$, $b > 0$ and $0 \leq p \leq 1$.

for the cases (a) $a < 1$, $b > 1$, (b) $a > 1$, $b > 1$, and (c) $a < 1$, $b < 1$,

its curve has a shape similar to the figures A, B and C respectively. The

distribution has the mean $E(p) = \frac{a}{a+b}$ and variance $\text{var}(p) = \frac{ab}{(a+b)^2(a+b+1)}$.

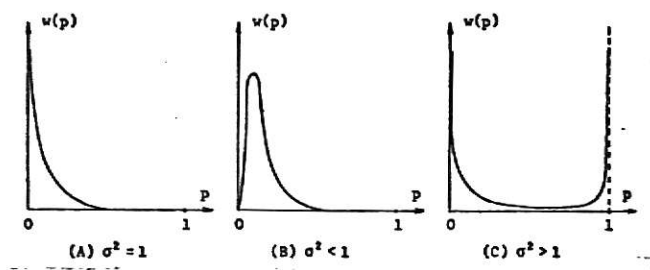


Fig. 8 The normal generated distribution. [12]

Now, we may ask which distribution fits actual data better. As the beta distribution involves the calculation of incomplete beta functions, with large values of b ; the numerical manipulation of the normal generated distribution is generally easier than that of beta distribution.

Let a single attribute sampling problem have large batches of size N and fraction defective p . Let the fraction defective p vary according to the normal generated prior density $w(p)$ given by equation 2.5.1. Let n be the sample size, c be the acceptance number and m be the observed number of defectives in the sample. Given n and p , we can say that m is a binomial variate. Hence we have,

$$E(m|p) = np \text{ and } E(m^2|p) = n(n-1)p^2 + np.$$

Hence the mean and variance of m can be derived as follows:

$$E(m) = nE(p) = n\bar{p},$$

$$\text{var}(m) = n(n-1)\sigma_p^2 + n\bar{p}\bar{q}, \text{ where } \bar{q} = 1-\bar{p}.$$

These are useful in the estimation of the mean \bar{p} and the variance σ_p^2 of the prior distribution.

Our main aim is to calculate the optimum values of n and c from a Bayesian viewpoint. Let p_r denote the breakeven quality, p_s denote the average sampling cost per item and p_m denote the average unavoidable minimum cost per item. The Bayesian single sampling plan, according to Hald (34), is given by

$$n = \lambda_1 \sqrt{N} + \lambda_2 \quad \text{and} \quad c = np_r + \beta_1 \quad (2.5.3)$$

where $\lambda_1^2 = \frac{rw}{2(p_s - p_m)}$,

$$\lambda_2 = \left\{ -3(1+4r)w + 4rr'w' + 9r^2w'' - \frac{12r^2(w')^2}{w} \right\} / (12rw)$$

and $\beta_1 = -r' - \frac{rw'}{w} - \frac{1}{2}$

where $w = w(p_r)$, $w' = w'(p_r)$, $w'' = w''(p_r)$, $r = p_r q_r$.

and $r' = 1 - 2p_r$.

$$\text{Now } p_m = p_r - \int_0^{p_r} (p_r - p)w(p)dp \quad (2.5.4)$$

The minimum regret is

$$R_0(N) = (2n - \lambda_1^2 - \lambda_2)(p_s - p_m) + O(n^{-1}) \quad (2.5.5)$$

Now λ_1^2 , λ_2 and β_1 can be rewritten as

$$\lambda_1^2 = p_r q_r \phi\left(\frac{m - \mu_r}{\sigma}\right) / \{2\sigma(p_s - p_m) \phi(\mu_r)\} \quad (2.5.6)$$

$$\lambda_2 = \{-3 - 12p_r q_r + 4v(1 - 2p_r) - 9u - 3v^2\} / (12p_r q_r) \quad (2.5.7)$$

and $\beta_1 = -1.5 + 2p_r - v \quad (2.5.8)$

where

$$u = p_r^2 q_r^2 \{ \mu_r (\mu_r - m) \sigma^{-2} + \sigma^{-2} - (1 + \mu_r^2) \} / \{ \phi(\mu_r) \}^2$$

and $v = p_r q_r \{ (\mu_r - m) \sigma^{-2} - \mu_r \} / \phi(\mu_r)$

For the normal generated prior distribution with the parameters m and σ^2 , we have

$$p_m = p_r \psi(-h) + \psi(-k) - \frac{1}{2\pi\sqrt{1-\rho^2}} \cdot \int_{-\infty}^{-h} \int_{-\infty}^{-k} \exp \left\{ -\frac{y^2 - 2\rho yz + z^2}{2(1-\rho^2)} \right\} dy dz \quad (2.5.9)$$

where

$$h = (m - \mu_r) / \sigma, \quad k = m / \sqrt{1 + \sigma^2}, \quad \rho = \sigma / \sqrt{1 + \sigma^2}$$

For example, consider the case

$$N = 3000, \quad \bar{p} = 0.06, \quad \sigma_p = 0.046, \quad p_r = 0.1, \quad p_s = 0.12, \quad \sigma^2 = 0.149.$$

Equations 2.5.6, 2.5.7, 2.5.8 and 2.5.9 yield,

$$p_m = 0.0531, \quad \lambda_2 = -22.56, \quad \lambda_1 = 0.69, \quad \beta_1^2 = 2.42 \text{ and } \lambda_1 = 1.56$$

Now the optimum value of n is

$$n = \lambda_1 \sqrt{N} + \lambda_2 = 63$$

and the optimum value of c is

$$c = np_r + \beta_1 = 7$$

The minimum regret is $R_0(N) = 9.8$

Hence we can conclude that a normal generated distribution may fit the actual data better than the widely used beta distribution as the normal generated distribution is derived from a realistic model. The curves from

these two distributions may have totally different shapes. The two distributions generally give somewhat similar optimum plans in normal circumstances, but in case of very large values of N and very small values of σ_p^2 , the plans may vary considerably.

ARTICLE #6

AN EMPIRICAL BAYES' ESTIMATOR FOR P (SUCCESS) IN THE BINOMIAL DISTRIBUTION BY B. S. GRIFFIN AND R. G. KRUTCHKOFF. [28]

2.6.1 Introduction

Here for an empirical Bayes situation, we will conduct a series of independent experiments. Let each experiment have its own parameter, say θ_m for the m -th experiment. Let each experiment possesses the following structure

- (i) Let the parameter θ be a function of ϕ which has unknown distribution function $G(\theta)$;
- (ii) Each experiment has its own value of θ .

Let us denote θ_1 for the first realization of ϕ and Z_1 for the observation from $p(Z_1|\theta_1)$; similarly for θ_2 and so on. For the current experiment m , θ_m is the parameter value and Z_m denotes the observed value from $p(Z_m|\theta_m)$. Let us denote $p(Z_m|\theta_m)$ by $B(n, \theta_m)$ assuming it is a binomial function. Hence Z_m is the number of successes that occurred in the m -th experiment. In this case, since the prior is not known, the Bayes' estimator can't be used. But from the data of past observations, a Bayes' estimator can be written. In short, the general principal underlying the empirical Bayes' approach consists of first determining the Bayes' estimator in a form which can be estimated and then using past information from the marginal distribution to estimate it.

2.6.2 Bayes' estimator

Let the loss for estimating θ when $\psi(Z)$ is the decision function be given by

$$L(\psi, \theta) = \frac{(\psi - \theta)^2}{\theta(1-\theta)} \quad (2.6.1)$$

The total risk of equation 2.6.1 is

$$R(\psi, G) = \int \sum_{Z=0}^n \frac{(\psi - \theta)^2}{\theta(1-\theta)} \binom{n}{Z} \theta^Z (1-\theta)^{n-Z} dG(\theta) \quad (2.6.2)$$

The above risk will be minimal if ψ is selected such that the posterior risk

$$r(\psi, G) = \int \frac{(\psi - \theta)^2}{\theta(1-\theta)} dG(\theta/Z) \quad (2.6.3)$$

is minimal for each value of Z .

Equation 2.6.3 can be differentiated with respect to ψ so as to get the minimal risk or Bayes' estimator

$$\psi_b(z) = \frac{\int \frac{\theta}{\theta(1-\theta)} dG(\theta/Z)}{\int \frac{1}{\theta(1-\theta)} dG(\theta/Z)} \quad (2.6.4)$$

By making substitution of $G(\theta)$ and the binomial loss function into 2.6.4, we get

$$\psi_b(Z) = \frac{\int \binom{n}{Z} \theta^Z (1-\theta)^{n-Z-1} dG(\theta)}{\int \binom{n}{Z} \theta^{Z-1} (1-\theta)^{n-Z-1} dG(\theta)}$$

which after multiplying with proper constants become

$$\psi_b(Z) = \frac{Z}{n-1} \cdot \frac{\int \binom{n-1}{Z} \theta^Z (1-\theta)^{n-Z-1} dG(\theta)}{\int \binom{n-2}{Z-1} \theta^{Z-1} (1-\theta)^{n-Z-1} dG(\theta)} \quad (2.6.5)$$

where $Z \neq 0, n$.

When the prior is $G(\theta)$, the marginal mass function of Z is given by

$$p^n(Z) = \int \binom{n}{Z} \theta^Z (1-\theta)^{n-Z} dG(\theta) \quad (2.6.6)$$

substituting 2.6.6 into 2.6.5, we get

$$\psi_b(Z) = \frac{Z}{n-1} \cdot \frac{p^{n-1}(Z)}{p^{n-2}(Z-1)}, \quad Z \neq 0, n \quad (2.6.7)$$

where $p^{n-1}(Z) = P[Z \text{ successes in sample of } (n-1)]$

and $p^{n-2}(Z-1) = P[(Z-1) \text{ successes in sample of } (n-2)]$

Both are marginal probabilities.

2.6.3 Empirical Bayes' estimator.

In this case, m independent observations are available from $p^n(Z)$.

Hence estimates of $p^{n-1}(Z)$ and $p^{n-2}(Z-1)$ can be constructed.

According to Lemon and Krutchkoff [42], we have

$$p_m^{n-1}(Z) = \left(\frac{1}{m}\right) \sum_{i=1}^m \binom{n-1}{Z} (\hat{\theta}_i)^Z (1-\hat{\theta}_i)^{n-Z-1}, \quad Z \neq 0, n \quad (2.6.8)$$

and

$$p_m^{n-2}(Z-1) = \left(\frac{1}{m}\right) \sum_{i=1}^m \binom{n-2}{Z-1} (\hat{\theta}_i)^{Z-1} (1-\hat{\theta}_i)^{n-Z-1}, \quad Z \neq 0, n \quad (2.6.9)$$

in order to estimate the required marginal probabilities. The term $\hat{\theta}_i$ is an estimator of the parameter for the i -th experiment. Let us employ the estimated Z_i/n for the $\hat{\theta}_i$'s in equations 2.6.8 and 2.6.9. Then substituting equations 2.6.8 and 2.6.9 back into 2.6.7, we have an empirical Bayes' estimator,

$$\psi_m = 0 \quad \text{if } Z = 0$$

$$\psi_m(Z) = \frac{Z \sum_{i=1}^m \binom{n-1}{Z} (Z_i/n)^Z (1-Z_i/n)^{n-Z-1}}{(n-1) \sum_{i=1}^m \binom{n-2}{Z-1} (Z_i/n)^{Z-1} (1-Z_i/n)^{n-Z-1}} \quad \text{if } Z \neq 0, n \quad (2.6.10)$$

$$= 1 \quad \text{if } Z = n$$

The empirical Bayes estimator $\psi_m(Z)$ given by equation 2.6.10 is a function of all previous observations on Z . Hence its risk is defined by

$$R_m(\psi_m, G) = \sum_{Z_1=0}^n \dots \sum_{Z_m=0}^n \int \sum_{y=0}^n L(\psi_m(y), \theta) p(y|\theta) dG(\theta) \prod_{i=1}^m p^n(Z_i)$$

For every different value of Z , ψ_m will be a different function. Hence for every sequence (Z_1, Z_2, \dots, Z_m) , there is one estimator. The risk R_m is the risk of $\psi_m(y)$ averaged over the total possible sequences. The risk of the maximum likelihood estimator $\psi_0(Z) = Z/n$, has a constant value $1/n$ irrespective of the value of $G(\theta)$.

The risk ratio is defined as

$$T_m = \frac{R_m(\psi_m, G)}{R(\psi_0, G)}$$

This risk ratio gives an indication of ψ_m 's performance. A simulation program is written to evaluate T_m with \bar{T}_m as T_m can not be evaluated theoretically.

T_m can be written as

$$T_m = (n/r) \sum_{j=1}^r (m(Z_{mj}) - \theta_{mj})^2 (\theta_{mj}(1-\theta_{mj}))^{-1}$$

where r = number of sequences (Z_1, Z_2, \dots, Z_m) generated

Z_{mj} = m -th observation in j -th sequence,

and θ_{mj} = θ -value generated for experiment m in sequence j .

In short, the simulation program has following steps:

- (1) a Pearson curve, $G(\theta)$, is created from the input values of the first four moments.
- (2) according to $G(\theta)$, a deviate, say θ_{mj} , is chosen at random.
- (3) from the binomial function $B(n, \theta_{mj})$, a binomial deviate Z_{mj} is chosen.
- (4) The estimate $\psi_m(Z_{mj})$ is calculated for fixed j .
- (5) the loss $L(\psi_m(Z_{mj}), \theta_{mj})$ is first calculated and then stored.
- (6) Steps 2,3,4,5 and 6 are repeated for m up through 100 and for $j = 1, 2, \dots, 1000$; this gives 1000 sequences of 100 observations each

- (7) an average is taken of the stored losses over the 1000 sequences for each value of m ;
- (8) for $m = 1, 2, \dots, 100$, \bar{T}_m and its standard error are calculated and then printed out.

It can be stated that in the beginning, as m increases, \bar{T}_m decreases sharply and afterwards as m further increases, \bar{T}_m decreases steadily.

ARTICLE #7

BAYESIAN SOLUTION OF THE SINGLE SAMPLE INSPECTION SCHEME BY G. B. WETHERILL [66].

In this article we assume that the prior distribution of a defective item is a mixed binomial distribution. According to Bernard [5], we have

$$(a_i, p_i), (i = 1, 2, \dots, k) \text{ where } \sum_{i=1}^k a_i = 1 \quad (2.7.1)$$

where a_i is the probability of having p_i defective items in a batch. We will assume that the loss incurred while accepting a batch when p_i is true is denoted by W_{1i} and the loss incurred while rejecting a batch when p_i is true is W_{2i} .

Let p_0 be the breakeven quality at which a batch can be accepted or rejected without loss. Let $p_1 > p_2 > \dots > p_k$. In this case W_{1j} will be zero for all j such that $p_j \leq p_0$ and W_{2j} will be zero for all j such that $p_j \geq p_0$.

If we assume p' and α as suitable constants, then the equations for an optimum solution can be put simply as

$$\frac{p_i}{q_i} = (p')^{i/\alpha} \quad (2.7.2)$$

2.7.1 Equation for the neutral line

The neutral line is defined as the locus of points (n, c_n) where c_n is the acceptance number, such that at these points the loss incurred while accepting a batch is equal to that incurred while rejecting it. Hence the equation of neutral line is

$$\sum_{i=1}^k a_i \binom{n}{c_n} p_i^{c_n} q_i^{n-c_n} (W_{1i} - W_{2i}) = 0 \quad (2.7.3)$$

substituting equation 2.7.2 in equation 2.7.3 and simplifying, we get

$$\sum_{i=1}^k a_i q_i^n (W_{1i} - W_{2i}) (p')^{c_n i/\alpha} = 0$$

This last equation can be rewritten as

$$\sum_{i=1}^k \lambda_i x^{i-1} = 0 \quad (2.7.4)$$

where $x = (p')^{c_n/\alpha}$

and $\lambda_i = a_i q_i^n (W_{1i} - W_{2i})$

If n is given, we can calculate the value of c with the help of this equation. Then after calculating x , we can have

$$c_n = \frac{\alpha \log x}{\log p'} \quad (2.7.5)$$

2.7.2 Equation for optimal sample size

The acceptance number c_n is determined by the neutral line once the sample size is known. Now our aim is to obtain an equation for the optimum value of n .

$R(n)$, the expected loss for a given n is give by

$$R(n) = n + \sum_{i=1}^k \sum_{r=0}^n P(r \text{ def.} | p_i \text{ true}) (\text{loss at } (n, r) | p_i \text{ true})$$

It is evident that for optimum value of n

$$R(n-1) > R(n) < R(n+1)$$

We will now consider the term $\{R(n+1) - R(n)\}$ i.e. difference between taking risk for sample size of $(n+1)$ and for sample size of n . At the optimum value of n , the term $\{R(n+1) - R(n)\}$ will change the sign from negative to positive.

Now by equating $\{R(n+1) - R(n)\}$ to zero, we will derive an expression for the optimum value of n .

We will assume that the slope of the neutral line is less than one. Hence $(n+1, c_n)$ is an acceptance point and $(n+1, c_n+1)$ is a rejection point. Consider the case where a batch is accepted at a sample size of n and rejected at a sample size of $(n+1)$. Hence the expression $\{R(n+1) - R(n)\}$ here has the value

$$\frac{1}{2} \sum_{i=1}^k a_i \binom{n}{c_n} p_i^{c_n} q_i^{n-c_n} p_i (w_{1i} - w_{2i})$$

In the above case the extra sample inspected is defective. Now consider the case where the extra sample inspected is effective. Let the batch be rejected at a sample size of n and accepted at a sample size of $(n+1)$. The expression $\{R(n+1) - R(n)\}$ becomes

$$\frac{1}{2} \sum_{i=1}^k a_i \binom{n}{c_n} p_i^{c_n} q_i^{n-c_n} q_i (w_{2i} - w_{1i})$$

Hence

$$R(n+1) - R(n) = \frac{1}{2} \sum_{i=1}^k a_i \binom{n}{c_n} p_i^{c_n} q_i^{n-c_n} (w_{2i} - w_{1i}) (2p_i - 1) + 1 \quad (2.7.6)$$

since $(p_i - q_i) = 2p_i - 1$

Now equating (2.7.6) to zero and using 2.7.2 and 2.7.4, we have

$$\sum_{i=1}^k \beta_i x^i = 1 / \binom{n}{c_n} \quad (2.7.7)$$

where $\beta_i = a_i p_i q_i^n (w_{1i} - w_{2i}) = p_i \lambda_i$.

Equations 2.7.4 and 2.7.7 essentially decide the optimum sampling plan for given prior probabilities and loss functions.

For $k = 2$, if $p_1 > p_0 > p_2$, equation 2.7.4 becomes

$$\frac{a_1 q_1^n w_{11}}{a_2 q_2^n w_{22}} = x \quad (2.7.8)$$

Hence from equation 2.7.5, we have

$$c_n = \frac{\alpha}{\log p} \left\{ \log \frac{a_1 w_{11}}{a_2 w_{22}} + n \log \left(\frac{q_1}{q_2} \right) \right\} \quad (2.7.9)$$

equation 2.7.7 can be written in terms of x as

$$a_1 p q_1^n w_{11} x - a_2 q_2^n p_2 w_{22} x^2 = 1 / \binom{n}{c_n}$$

Hence from equation 2.7.8, we have

$$\frac{a_{22} w_{22}}{(a_1 w_{11})^2 (p_1 - p_2)} \left(\frac{q_2}{q_1} \right)^n = \binom{n}{c_n} \quad (2.7.10)$$

Hence

$$\log \left\{ \frac{a_2 W_{22}}{(a_1 W_{11})^2 (p_1 - p_2)} \right\} + n \log \left(\frac{q_2}{q_1} \right) = \log \binom{n}{c_n} \quad (2.7.11)$$

We can write the above equation as

$$y(n) = \log \left\{ \frac{a_2 W_{22}}{(a_1 W_{11})^2 (p_1 - p_2)} \right\} + n \log \left(\frac{q_2}{q_1} \right) \quad (2.7.12)$$

$$\text{and } Z_n = \log \binom{n}{c_n} \quad (2.7.12)$$

The optimum value of n is that one at which $y(n) = z(n)$

2.7.3 Illustrative example

Suppose we have $p_1 = 0.09132$, $p_2 = 0.01$, $\alpha = 1$, $p' = 0.01$,
 $W_{11} = W_{22} = 2000$, $a_1 = 0.20$, $a_2 = 0.80$

The equation of the neutral line is

$$c_n = 0.6034 + 0.0373n$$

and the equation for the optimum sample size is

$$\log \binom{n}{c_n} = -0.9102 + 0.0788n$$

The following values for $n, c_n, y(n)$ and $Z(n)$ were obtained

n	C_n	$y(n)$	$Z(n)$
30	1.7225	1.4343	1.4771
70	3.2148	4.6069	4.9986
100	4.3339	6.9714	7.0281
105	4.5204	7.3655	7.3689
106	4.5578	7.4443	7.4390
110	4.7070	7.7595	7.7095

At $n = 105$, the value of $y(n)$ and $Z(n)$ are almost equal. Hence our optimum sampling plan is $n = 105$ and $C_n = 4$. This is a relatively rapid technique for computing n and C_n .

ARTICLE #8

BAYESIAN ESTIMATION FOR THE PROPORTIONS IN A MIXTURE OF DISTRIBUTIONS

BY J. BEHBOODIAN [3].

Let us consider the equation

$$m(x) = pf(x) + qg(x) \quad (2.8.1)$$

where $f(x)$ and $g(x)$ are two known, independent probability density functions with $0 < p < 1$, and $q = 1-p$. As $f(x)$ and $g(x)$ are independent from each other, two different values of p will give two different values of $m(x)$. Boes [7] has already investigated the estimation problem of proportions in a finite mixture of distributions. Here we will consider the information about the proportion p prior to taking a sample. A prior distribution for p generally expresses such information. By applying Bayes' theorem, the prior distribution yields a posterior distribution of p .

Since the posterior distribution is the product of the likelihood function and prior density, we will first express its likelihood function in a suitable form before introducing a prior distribution for p .

2.8.1 The likelihood function

Let the probability density function, equation 2.8.1, have a population X_1, X_2, \dots, X_n as a random sample. Now the likelihood function of p for the values x_1, x_2, \dots, x_n as the random sample is

$$L(p) = \prod_{i=1}^n [pf(x_i) + qg(x_i)] \quad (2.8.2)$$

where $0 < p < 1$

We can eliminate non-informative x_i 's from the observed sample for which $f(x_i) = g(x_i)$. This case may happen in case of a mixture of two uniform distributions defined on two different overlapping intervals of equal length. But as n becomes large, the chance of such event becomes small due to distinction of $f(x)$ and $g(x)$. Therefore we will assume that $f(x_i) \neq g(x_i)$ for all x_i 's to avoid exceptional cases.

We can expand the right hand side of equation 2.8.2 to write $L(p)$ in a convenient form. But for convenience, let us express the right hand side of equation 2.8.2 by $h(x_1, x_2, \dots, x_n)$. Using conditional density, we have

$$h(x_1, x_2, \dots, x_n) = \sum_{k=0}^n h(x_1, x_2, \dots, x_n | E_k) P(E_k) \quad (2.8.3) \quad *$$

where E_k is the event that exactly k of the x_i 's have density $f(x)$ and others have density $g(x)$. Hence how

$$P(E_k) = \binom{n}{k} p^k q^{n-k} \quad (2.8.4)$$

The joint density of x_1, x_2, \dots, x_n is

$$h_{t_k}(x_1, x_2, \dots, x_n) = \prod_{A_k} f(x_a) \prod_{B_k} g(x_b) \quad (2.8.5)$$

where t_k is a partition of the $\{1, 2, \dots, n\}$ into two sets A_k and B_k with k elements into A_k . Let T_k express the set of all such partitions and once more using the conditional density, we have

$$S_k(x_1, x_2, \dots, x_n) = h(x_1, x_2, \dots, x_n | E_k) = \sum_{t_k \in T_k} h_{t_k}(x_1, x_2, \dots, x_n) / \binom{n}{k} \quad (2.8.6)$$

where $s_k(x_1, x_2, \dots, x_n)$ is a symmetric n -variate density. Now from equations 2.8.3, 2.8.4 and 2.8.6, we have

$$L(p) = h(x_1, x_2, \dots, x_n) = \sum_{k=0}^n \binom{n}{k} p^k q^{n-k} s_k(x_1, x_2, \dots, x_n) \quad (2.8.7)$$

Hence, the joint density of the random sample x_1, x_2, \dots, x_n is a binomial mixture of the densities $s_k(x_1, x_2, \dots, x_n)$ expressed by equations 2.8.5 and 2.8.6.

2.8.2 Bayesian estimation for p .

Let the random variable p be expressed by a beta density function

$$B(p; u, v) = \frac{\Gamma(u+v)}{\Gamma(u) \Gamma(v)} p^{u-1} (1-p)^{v-1} \quad (2.8.8)$$

where $0 < p < 1$ and $u > 0, v > 0$. Now the posterior density of p is given by

$$\pi(p | x_1, x_2, \dots, x_n) \propto L(p) \times B(p; u, v) \quad (2.8.9)$$

where \propto denotes proportionality. Omitting the multiplier of equation 2.8.8 and the using equations 2.8.7 and 2.8.8, we get

$$\pi(p | x_1, x_2, \dots, x_n) \propto \sum_{k=0}^n \binom{n}{k} s_k(x_1, x_2, \dots, x_n) p^{k+u-1} q^{n-k+v-1} \quad (2.8.10)$$

The constant of proportionality can be easily found from the fact that the posterior density must integrate to one. It can be shown that

$$\pi(p | x_1, x_2, \dots, x_n) = \sum_{k=0}^n w_k B(p; k+u, n-k+v) \quad (2.8.11)$$

w_k can be expressed as

$$w_k = \gamma_k s_k(x_1, x_2, \dots, x_n) / \sum_{j=0}^n \gamma_j s_j(x_1, x_2, \dots, x_n) \quad (2.8.12)$$

$$\text{where } \gamma_k = \binom{n}{k} / \binom{n+u+v-1}{k+u-1} \quad (2.8.13)$$

The w_k depends upon the observed samples and parameters of the prior; the contribution of the population is reflected in $s_k(x_1, x_2, \dots, x_n)$ and the contribution of the prior is reflected in γ_k .

By using equation 2.8.11 and the mean and variance of a beta distribution, the posterior mean and posterior variance of p can be shown as,

$$E_{\Pi}(p) = \sum_{k=0}^n w_k \frac{k+u}{n+u+v} \quad (2.8.14)$$

$$\text{var}_{\Pi}(p) = \sum_{k=0}^n w_k \frac{(k+u)(n-k+v)}{(n+u+v)^2(n+u+v+1)} + \sum_{k=0}^n w_k \left(\frac{k+u}{n+u+v} - E_{\Pi}(p) \right)^2$$

By using a family of finite mixtures of beta distributions, we can get a diverse variety of distributions for expressing knowledge about p . If we take a member of a family for the prior density of p , then it is noted that the posterior density will be a member of the same family. Hence a family of finite mixtures of beta distributions is conjugate with regard to the likelihood function.

CHAPTER 3

SHORT SUMMARIES OF IMPORTANT ARTICLES

(3.1) Rubin, H. and Sethuraman, J., "Bayes risk efficiency." [59]

In this paper, efficiency of tests are considered from a Bayesian point of view. For two test procedures the ratio of the sample sizes needed to obtain equal expected risks is defined to be the Bayes risk efficiency (BRE). This efficiency is relatively insensitive to the weight functions. The discussion presented also gives the optimum choice of the significance level and consequences of a non-optimum choice. Bayes expected risks obtained by varying the critical region based on one or a few test statistics are approximated for large samples. BRE extends easily to the comparison of such test procedures as the Kolmogorov-Smirnov and Anderson-Darling tests, and to multiparameter testing problems. The problem of BRE is treated here for a one sided hypothesis.

(3.2) Cacoullos, T. "Comparing Mahalanobis distances II: Bayes procedures when the mean vectors are unknown". [11]

Let II_i be multivariate normal populations with means μ_i , $i = 0, 1, \dots, k$ ($k \geq 2$), respectively, and with the same known covariance matrix Σ . The problem of selecting the nearest II_j , $j = 1, \dots, k$, to II_0 is considered when only μ_0 is known and μ_1, \dots, μ_k are estimated from samples of equal size out of each II_j . A unique Bayes procedure is obtained by restricting attention to decision rules which are invariant under a subgroup of affine transformations on $k p$ -space and the symmetric group of permutations on the sample means $\bar{x}_1, \dots, \bar{x}_k$. The theory developed, which is actually of wider applicability, is further applied to the problem of selecting, among k pairs of normal populations, a pair

with minimum distance between its members. The corresponding problems when the common covariance matrix is also unknown or all the mean vectors are unknown are also discussed.

- (3.3) Van Ryzin, J., "Bayes risk consistency of classification procedures using density estimation". [64]

This paper introduces the concept of Bayes risk consistency for classification procedures. This notion allows to examine asymptotic properties of classification procedures which are based on methods of general (non-parametric) density estimation and related results. The advantage of these procedures is that they do not require assumptions on the exact parametric form of the distributions, while the advantage of the concept of Bayes risk consistency is that it aids considerably in the selection of estimating procedures under limited knowledge of the class distributions.

- (3.4) Kappenman, R. F., Geisser S. and Antle, C. E., "Bayesian and Fiducial solutions for the Fieller-Creasy problem". [39]

In biological assay work where one is interested in the relative potency of two drugs or treatments, the following problem arises. Suppose a random sample is available from a bivariate normal population with vector μ , where $\mu = (\mu_1, \mu_2)$. We can find the range of values which can be ascribed to $\eta = \mu_2/\mu_1$, with any desired degree of probability. In this paper the results of an investigation of the Bayesian theory of inference as it applies to the problem are presented.

(3.5) Geisser, S., "Bayesian analysis of growth curve", [24]

From a Bayesian viewpoint we initiate the study of the generalized growth curve model $E(Y_{p \times N}) = X_p \times_m \tau_m \times A_{r \times N}$ where X and A are known matrices of rank $m < p$ and $r < N$ respectively; τ is unknown and the columns of Y are independent p -dimensional multinormal variates having unknown variance matrix Σ . A Bayesian justification is presented for Rao's adjusted estimator τ of τ' , the set of unknown parameters, as well as an estimating region. The order unadjusted estimator of τ is also shown to obtain either for a structured covariance matrix or for an augmented location model. The problem of estimating regions for future observations from this model, given the past sample, is discussed in detail, both for an arbitrary and a structured covariance matrix.

(3.6) Novick, M. R. and Grizzle, J. E., "A Bayesian approach to the analysis of data from clinical trials", [51]

A Bayesian logical probability approach is described and its advantages with respect to constructing prior distributions for Bayesian analysis are discussed. Some standard Bayesian distribution theory for categorized data is summarized. Data from an on-going experiment to compare the relative efficiency of four operative treatments for ulcer are used both to pinpoint some of the practical problems involved in clinical trials and to demonstrate the usefulness of Bayesian inference methods when facing these problems.

- (3.7) Ando, A. and Kaufman, G. M., "Bayesian analysis of the independent multinormal process - neither mean nor precision known", [1]

Here under the assumption that neither the mean vector nor the variance-covariance matrix are known with certainty, the natural conjugate family of prior densities for multivariate normal process is identified. Prior- posterior and preposterior analysis is done assuming that the prior is in the natural conjugate family. Here a procedure is presented for obtaining non-degenerate joint posterior and preposterior distributions of all parameters even when the number of objective sample observations is less than the number of parameters of the process.

- (3.8) Antelman, G. R., "Insensitivity to non- optimal design in Bayesian decision theory", [2]

Two simple inequalities involving two parameters, expected terminal losses, expected sampling losses, and optimal (Bayes) and non-optimal sample sizes are shown to hold for several fixed sample size decision problems. The two parameters depend on the type of loss structure assumed. One inequality relates to the division of total expected losses for a sample of optimal size between expected terminal losses and expected sampling losses. The other inequality gives upper bounds on the ratio of total expected losses at non-optimal sample sizes to those at the optimal sample size. The latter inequality shows that total expected losses are often quite insensitive to the use of non-optimal sample sizes; in conjunction with optimal sample size formulas; it can be used to show that total expected losses are also insensitive to the use of a 'wrong' prior distribution or the wrong cost parameters.

The inequalities are shown to hold for several two-action problems on the mean of a Normal process, ten quadratic loss estimation problems involving Normal, Bernoulli, and Poisson processes, and one linear loss estimation problem on the mean of a Normal process, in each of these problems, a conjugate prior distribution is assumed.

(3.9) Geisser, S., "A Bayes approach for combining correlated estimates",
[23]

A Bayes solution is supplied for an estimation problem involving a sample from a multivariate normal population having an arbitrary unknown covariance matrix, but a vector mean whose components are all equal. Assuming that a particular unnormed prior density is a convenient expression for displaying prior ignorance, it is then demonstrated that a posterior interval for this common mean can be based on student's t distribution. If prior information can be conveniently represented by a natural conjugate prior density, the posterior interval will also depend on student's t . An extension is made to the case of estimating the constant difference between two parallel profiles.

(3.10) Novick, M. R. and Hall, W. J., "A Bayesian indifference procedure",
[52]

In a logical probability approach to inference, distributions on a parameter space are interpretable as representing states of knowledge, and any prevailing state of knowledge may be taken to have been arrived at from a previous state of ignorance followed by an accumulation of prior data. In this paper an indifference procedure is introduced that requires postulating what size and what kind of samples will and will

not permit statistical inference and prediction - e.g., one observation from a two-parameter normal model is not sufficient to permit inference about the variance but two observations are. In essence, the procedure stipulates that prior indifference distributions be improper but become proper after an appropriate minimal sample. With some limitation on the family of priors considered, this procedure permits unique specification of indifference for the more commonly encountered statistical models. Furthermore, these specifications are affected neither by change of the scale of measurement of the observations, nor by the sampling rule.

- (3.11) Bhattacharya, S. K., "Bayesian approach to life testing and reliability estimation", [6]

Bayesian analysis of the exponential model, based on life tests that are terminated at preassigned time points or after preassigned number of failures, has been developed. For the prior distribution of the parameter involved, uniform, inverted gamma and exponential densities have been examined. The estimation of the reliability function has also been carried out by using Bayesian methods and the case of 'attribute testing' has been considered briefly. The role of prior quasi-densities when a life tester has no prior information has been illustrated and it has been observed that the reliability estimate for a diffuse prior which is uniform over the entire positive real line closely resembles the classical MVU estimate obtained by Pugh.

(3.12) Winkler, R. L., "The assessment of prior distributions in Bayesian analysis", [67]

In the Bayesian framework, quantified judgements about uncertainty are an indispensable input to methods of statistical inference and decision. Ultimately, all components of the formal mathematical models underlying inferential procedures represent quantified judgements. In this study, the focus is on just one component, the prior distribution, and on some of the problems of assessment that arise when a person tries to express prior distributions in quantitative form. The objective is to point toward assessment procedures that can actually be used.

One particular type of statistical problem is considered and several techniques of assessment are presented, together with the necessary instruction so that these techniques can be understood and applied. A questionnaire is developed and used in a study in which people actually assess prior distributions. The results indicate that, by and large, it is feasible to question people about subjective prior probability distributions, although this depends on the assessor and on the assessment techniques used. A revised questionnaire, which is aimed at potential users of the assessment procedures and future investigators in the area of probability assessment, is presented.

(3.13) Box, G. E. P. and Tiao, G. C., "Bayesian estimation of means for the random effect model", [8]

The problem of estimating the means in the one-way random effect model $y_{jk} = \theta_j + e_{jk}$ is considered from a Bayesian viewpoint. Posterior distributions of the θ_j are obtained under the assumption that the θ_j

are independently drawn from a normal population $N(\theta, \sigma_1^2)$ and that the e_{jk} are independent random errors having a $N(0, \sigma_1^2)$ distribution. It is shown that the posterior distributions of the θ_j are clustered more closely together than are the corresponding distributions for a fixed effect model. A numerical example is given.

- (3.14) Hill, B. M., "Posterior distribution of percentiles: Bayes' theorem for sampling from a population", [36]

A Bayesian approach to inference about the percentiles and other characteristics of a finite population is proposed. The approach doesn't depend upon the use of parametric models. Some related questions concerning the existence of exchangeable distributions are considered. It is shown that there are no countably additive exchangeable distributions on the space of observations which give ties probability 0 and for which a next observation is conditionally equally likely to fall in any of the open intervals between successive order statistics of a given sample.

- (3.15) Lee, T. C., Judge, G. G. and Zellner, A., "Maximum likelihood and Bayesian estimation of transition probabilities", [41]

In this paper, maximum likelihood and Bayesian methods are presented for estimating transition probabilities when data in the form of aggregated proportions are available. The probability function for the observed proportions is assumed to have a multinomial distribution under the Lexis scheme. The multivariate beta distribution is used as the prior probability density function in formulating the Bayesian estimator. The results of some Monte Carlo experiments provide some evidence on the sampling properties of several alternative estimators.

- (3.16) Zellner, A. and Tiao, G. C., "Bayesian analysis of the regression model with autocorrelated errors", [68]

In this paper-regression models with error terms generated by a first order autoregressive scheme are analyzed from a Bayesian point of view. Methods are developed for computing posterior distributions of regression coefficients and the parameter of the autoregressive process. The relationship of this approach to sampling theory approaches is briefly discussed.

- (3.17) Hoadley, B., "The compound multinomial distribution and Bayesian analysis of categorical data from finite populations", [37]

A Bayesian analysis of the parameter vector, $W(W_j = \text{number of elements in category } j)$, of a multivariate hypergeometric distribution is considered. It is shown that if, a priori, W is compound multinomial then a posteriori W is a translated compound multinomial. Many properties of the compound multinomial distribution are derived. These include joint moments of all orders; a characterization in terms of independent compound poisson variables, conditional distribution of one subvector given another, and joint distributions of disjoint and overlapping sums of the component.

It is shown that in applications related to analysis of variance and contingency tables, the parameters of interest are functions of W . A method is proposed for approximating posterior distributions of such parameters, and a numerical example involving a performance index defined on a contingency table is presented.

- (3.18) Waller, R. A. and Duncan, D. B., "A Bayes rule for the symmetric multiple comparisons problem", [65]

A simple LSD (least significant difference) rule is presented for simultaneously testing the difference between n treatments considered in all possible pairs. This rule is based on the same multiple decision theory model except for a modified and extended use of a conjugate chi-square density in the prior. The new rule has the same intuitively appealing dependence on the between-treatment F ratio, varying from a sensitive comparisonwise- α -like rule when F is large or moderate, to a conservative experimentwise- α -like rule when F is small.

- (3.19) Hoadley, B., "A Bayesian look at inverse linear regression", [38]

The model considered in this paper is simple linear regression ($Ey_i = \beta_1 + \beta_2 x_i$, $i = 1, \dots, n$), and the problem is to make statistical inferences about an unknown value of x corresponding to one or more additional observed values of y .

- (3.20) Girshick, M. A. and Rubin, H., "A Bayes approach to a quality control model", [25]

This paper deals with a class of statistical quality control procedures and continuous inspection procedures which are optimum for a specified income function and a production model which can only be in one of four states, two of which are states of repair, with known transition probabilities. The Markov process, generated by the model and the class of decision procedures, approaches a limiting distribution and the integral equations from which the optimum procedures can be derived are given.

(3.21) Lindley, D. V., "The Bayesian analysis of contingency tables", [45]

This paper describes how data from a multinomial distribution, and in particular data in the form of a contingency table, may be studied by using a prior distribution of the parameters and expressing the results in the form of a posterior distribution of the parameters. The analysis used must depend on the prior distribution and the forms described here only applies to a certain type of prior knowledge but it is believed that this type is of frequent occurrence. The binomial situation is first considered and the results obtained there suggest a general result for the multinomial distribution, which is then established. A few remarks on Bayesian analysis in general enable the result to be applied, first to certain multinomial problems and then, with the aid of another general result, to contingency tables. The method used there has close connections with the analysis of variance and these connections are examined, particularly with a view to simplifying the analysis of contingency tables involving three or more factors.

(3.22) Samuel, E., "An empirical Bayes approach to the testing of certain parametric hypotheses", [60]

In this paper, initially the empirical Bayes approach is described. Optimal empirical Bayes rules are given for the problem of testing a simple hypotheses against a simple alternative. A limit theorem is proved which is used to obtain optimal empirical Bayes rules for testing one and two sided hypotheses about the parameters in the Poisson, geometric, negative binomial and binomial distributions. The same methods are used to obtain optimal empirical Bayes rules for testing

hypotheses about parameters in continuous distributions of the exponential family. Examples of areas of applications are given and applications of the above methods in the compound decision problem are discussed.

- (3.23) Draper, N. R. and Guttman, I., "Unequal group variances in the fixed-effects one-way analysis of variance: A Bayesian side-light", [17]

In two comprehensive papers, Box (1954), *Ann. Math. Stat.*, 25, 290-302 & 484-498, examined the effect of inequality of variance on the standard F-tests for one-way and two-way analysis of variance (ANOVA) classifications. These papers provide a number of theorems on the exact and approximate distributions of various quadratic forms and ratios of quadratic forms, and applied them to the ANOVA situations mentioned above. In the present paper certain results found by Box for the one way classification are examined, and how a Bayesian analysis can throw light on the consequences of Box's work in a given experimental situation is also observed.

- (3.24) Maritz, J. S., "Smooth empirical Bayes estimation for one-parameter discrete distributions", [46]

A study is made of the simple empirical Bayes estimators proposed by Robbins (1956). (*Proc. 3rd Berkeley Symposium on Math. Stat. and Prob.*). These estimators are compared with best conventional estimators in terms of their expected squared error loss. The object of the study is to determine the amount of prior data which would be needed for the empirical Bayes estimator to be preferred to the conventional estimator.

It is concluded that the required number of previous observations is likely to be too large for the simple empirical Bayes estimators to be useful in practice. Smooth empirical Bayes estimators are proposed, which make more effective use of prior results. A method of smoothing is developed which is based on estimating a step function approximation to the prior distribution of the parameter. Some examples are studied in detail, and the results indicate that these smooth empirical Bayes estimators are potentially useful in practice.

- (3.25) Springer, M. D. and Thompson, W. E., "Bayesian confidence limits for the product of N binomial parameters", [62]

The posterior probability density function of the product of N binomial parameters is derived in closed form using the Mellin integral transform. This result permits the computation of Bayesian confidence limits for the product of an arbitrary number of binomial parameters, and has an immediate application to problems of reliability and test of independent cascaded subsystems.

- (3.26) Rutherford, J. R. and Krutchkoff, R. G., "The empirical Bayes approach: estimating the prior distribution", [57]

There is a random variable A distributed according to a specific but unknown prior distribution G from an appropriate class G_p . The random variable $A = \lambda$ is unobservable but another random variable $X = x$, distributed with known conditional distribution function $F(x/\lambda)$, is observable. We construct estimators $G_n(\lambda)$ of $G(\lambda)$ such that $\lim E[\{G_n(\lambda) - G(\lambda)\}^2] = 0$ and we use $G_n(\lambda)$ to estimate the posterior

distribution $G(\lambda/x)$ and hence to construct consistent estimators of posterior confidence intervals.

- (3.27) Maritz, J. S., "Smooth empirical Bayes estimation for continuous distributions", [47]

Although the conditions for application of the technique of empirical Bayes estimation of the parameter of a distribution may exist, the form of the prior distribution is generally unknown. This often creates major difficulties in the determination of empirical Bayes estimators. The simple device of approximating the prior distribution by a step function is used to overcome this problem and to obtain smooth empirical Bayes estimators. This paper supplements earlier work by considering continuous distributions and multiple past and current observations. The effectiveness of the proposed smooth empirical Bayes estimators is examined, and the results indicate that they can be substantially better than optimum non-Bayes estimators, even when the amount of past data is small.

- (3.28) Krutchkoff, R. G., "A supplementary sample non-parametric empirical Bayes approach to some statistical decision problems", [40]

When an estimating problem is routine, it is often possible to consider the parameter being estimated as a random variable. The data obtained to estimate previous values of parameter then contain information which can be used to advantage in estimating the present parameter. Besides this data it is assumed that there are supplementary estimates

of the previous parameters, perhaps in the form of customer feedback. All the probability distributions are assumed to be unknown. The estimating procedure given here is shown to be asymptotically optimal, and by a Monte Carlo example to have good small sample properties.

- (3.29) Maritz, J. S., "On the smooth empirical Bayes approach to testing of hypotheses and the compound decision problem", [48]

The smooth empirical Bayes approach is based on using past observations to estimate a specified type of approximation to the prior distribution. This approach is applied to problems of empirical Bayes hypothesis testing and to the compound decision problem. Problems involving composite hypotheses are considered which have received much less attention than the case of two simple hypotheses. The methods are shown to be asymptotically optimal and results of studies involving finite sample numbers are reported, giving some indication of the rate of approach to optimality in specific cases.

- (3.30) Draper, N. R. and Guttman, I., "Some Bayesian stratified two-phase sampling results", [18]

In this paper some results concerning the optimum allocation of sampling effort among k - strata at the second phase of a two-phase sampling procedure are derived by using information obtained from the first phase. Two different approaches are used; a Bayesian posterior analysis and a Bayesian preposterior analysis. Two different allocation methods are derived and illustrated with some numerical examples, for cases where some or all of the nuisance parameters are unknown.

- (3.31) Rutherford, J. R. and Krutchkoff, R. G., "Some empirical Bayes techniques in point estimation", [58]

In point estimation with a squared-error loss function the Bayes estimator is the posterior mean. In the empirical Bayes approach we must construct a consistency sequence of estimators for this posterior mean using past experience. This construction is done here for four general families of conditional distributions which include as special cases: the Poisson, the gamma, the normal and the uniform.

- (3.32) Maritz, J. S., "Empirical Bayes estimation for the Poisson distribution", [49]

A number of established methods of empirical Bayes estimation for the Poisson distribution are surveyed, and their performances are compared. A new method of smooth empirical Bayes estimation is presented, and its performance is also studied. The question of assessing the effectiveness of the empirical Bayes approach in practice is examined.

- (3.33) Griffin, B. S. and Krutchkoff, R. G., "Optimal linear estimators: an empirical Bayes version with application to the binomial distribution", [29]

An empirical Bayes estimator is one which estimates the posterior mean by making use of past data. For certain conditional distributions no empirical Bayes estimator can be found which converges to the posterior mean as past data are accumulated. However, an optimal linear estimator for a parameter, say θ , can often be found. This optimal linear estimator depends upon the first two prior moments, both of which can often be

estimated. The resulting estimator has been simulated under the assumption that the conditional distribution is binomial and these simulations have shown its risk substantially smaller than the risk of the maximum likelihood estimator.

- (3.34) Bennett, G. K. and Martz, H. F., "A continuous empirical Bayes smoothing technique", [4]

Discrete empirical Bayes smoothing techniques essentially attempt to approximate the prior distribution function. Here a continuous smoothing technique which is based on a smooth and continuous approximation to the prior density function is presented. Results from a Monte Carlo study of the Poisson distribution are reported which show that the continuous smoothing technique has desirable small sample properties. Some comparisons with discrete smoothing techniques are also made.

- (3.35) Leonard, T., "Bayesian methods for binomial data", [43]

A Bayesian procedure is obtained for the simultaneous estimation of the parameters of m binomial distributions. The method uses logistic transformations for the parameters and an exchangeable prior distribution. Information is combined between the binomial distributions to obtain estimates which, under certain circumstances, will be superior to the usual proportions. This paper is intended as a forerunner to a more general theory for the analysis of nonlinear models.

- (3.36) El-Sayyad, G. M. and Freeman, P. R., "Bayesian sequential estimation of a Poisson process rate", [20]

This paper provides numerical and analytical solutions to the problem of estimating the rate of a poisson process. Optimal designs are obtained for various loss functions and the method of analysis is valid for any other loss function. The cost of sampling plays a fundamental rule and since there are many practical situations where there is a time cost and an event cost, a sampling cost per observed event and a cost per unit time are both included.

- (3.37) Leonard, T., "A Bayesian method for histograms", [44]

This paper describes a Bayesian procedure for the simultaneous estimation of the probabilities in a histogram. A two-stage prior distribution is constructed which assumes that probabilities corresponding to adjacent intervals are likely to be closely related. The method employs multivariate logit transformations, and a covariance structure similar to that assumed in the first-order auto regressive process. Posterior estimates are obtained which combine information between the intervals and have the practical effect of smoothing the histogram.

- (3.38) Martz, H. F. and Lian, M. G., "Empirical Bayes estimation of the binomial parameter", [50]

A smooth empirical Bayes estimator for a binomial parameter is derived. The risk in using this estimator is compared by simulation with that of eight other binomial estimators. For these estimators not having a simple closed form risk, suitable regression equations on the simulation

parameters are obtained. These equations are used to identify the regions of superiority of each of the estimators. A numerical example is provided.

- (3.39) Gunel, E. and Dickey, J., "Bayes factors for independence in contingency tables", [31]

The null hypothesis of row-column independence in a two-way contingency table can be expressed as a constraint on the parameters in various standard statistical sampling models. Four sampling models are considered, which are related by nested conditioning. By having the prior distribution in any one model induce the prior distribution in each further conditioned model, it is shown that the Bayes factors for independence will factorize, and thereby expose the evidence residing in the marginal row and column of the table. Bounds on the marginal Bayes factors justify, in a weak sense, Fisher's practice of conditioning. A general theorem is given for factorized Bayes factors from a factorized likelihood function.

- (3.40) Rao, J. N. K, and Ghangurde, P. D., "Bayesian optimization in sampling finite populations", [55]

The problem of optimum allocation in sampling finite populations using prior information is considered. The following cases are investigated: (1) stratified simple random sampling with known strata sizes; (2) Neyman's double sampling with unknown strata sizes; (3) the Hansen-Hurwitz method for the non-response problem; (4) Two stage random sampling. The optimum allocation in each case is obtaining by minimizing the expected posterior variance of the mean subject to constraints. The results are extended

to multiple prior distributions and/or multiple characters. The solutions are distribution-free and also free from the assumption of infinite populations and/or known variances. Attention is given to "data-based" prior distributions.

- (3.41) Romberg, H. F., "Continuous sequential testing of a Poisson process to minimize the Bayes risk", [56]

A method of testing constant failure rates is presented, where the distribution of failure times is assumed to be exponential. The failure rate is unknown but two alternative failure rates are hypothesized. The objective of the procedure is to minimize the decision cost, and the observational costs accumulated during the test. A Bayes procedure is found whereby the prior probability that one of the failure rates is true is updated until it reaches a previously determined decision point. The exact solutions and approximate solutions for the decision points for the minimum Bayes Risk are provided for the general renewal case.

- (3.42) Dayananda, R. A. and Evans, I. G., "Bayesian acceptance - sampling schemes for two sided tests of the mean of a normal distribution of known variance", [15]

This article is concerned with the problem of deciding whether the mean θ of a normal distribution of known variance lies in a specified finite interval $(\bar{\theta}, \theta^+)$. Consideration is given to prior information on θ and to quadratic and piecewise linear utility structures. Computer aided methods are described for obtaining the optimum decision rule given a sample of observations and for obtaining the optimal sample size when the sampling cost is a linear function of the sample size. Some simpler approximate methods are also described.

- (3.43) Good, I. J. and Crook, J. F. (1974), "The Bayes/non-Bayes compromise and the multinomial distribution", Am. Stat. Ass. J., 69, 711-720, [26]

Compromises between Bayesian and non-Bayesian significance testing are exemplified by examining distributions of criteria for multinomial equiprobability. They include Pearson's χ^2 , the likelihood ratio, the Bayes factor F , and a statistic G that previously arose from a Bayesian model by "Type II maximum likelihood." Its asymptotic distribution, implied by the theory of the "Type II likelihood ratio" is remarkably accurate into the extreme tail. F too can be treated as a non-Bayesian criterion and is almost equivalent to G . The relationship between F and its own tail area sheds further light on the relationship between Bayesian and "Fisherian" significance.

- (3.44) Duncan, D. B., "A Bayesian approach to multiple comparisons", [19]

The main concern of this paper is the symmetric multiple comparison problem of simultaneously testing the differences between several sample means taken in all possible pairs. Past and present procedures for solving the problem are discussed and illustrated with emphasis on their wide disagreements. A multiple-decision approach to the problem with a simple Jeffreys-like Bayesian emphasis is developed in depth. A working minimum-average-risk procedure is derived which is found to have much the same form as the simple Fisher LSD (least significant difference) rule, but with the LSD determined as a specifically defined function of the between-treatments observed F ratio. The overall analysis is of more general interest. It provides a practical example of potentially fruitful uses of decision-theoretic Bayesian techniques in a larger class of similar important experimental inference problems.

- (3.45) Bracken, J., "Percentage points of the beta distribution for use in Bayesian analysis of Bernoulli processes", [9]

Here percentage points of the beta distribution Z_{II} are tabulated such that $F_{\beta}(Z_{II}/r, n) = II$. The percentage points are given to four decimal places for $II = .01, .05, .95, .99$, $n = 2(1)30, 40, 50$ and $r = 1(1)[n/2]$. The tables are primarily for use in Bayesian analysis of Bernoulli processes. Numerical integration of a function with respect to the beta distribution is also discussed.

- (3.46) Grosh, D. L., "A Bayes sampling allocation scheme for stratified finite populations with hyperbinomial prior distributions", [30]

When sampling is carried out independently for the k - strata of a finite stratified dichotomous population (defective vs. standard items), and the number x_i of defectives per stratum sample is observed, the corresponding probability function for $X = (x_1, \dots, x_k)$ is the product of hypergeometric functions which depend on the sample size n_i , the stratum sizes N_i , and the number of defectives m_i in the stratum ($i = 1, \dots, k$). It is assumed that prior information is available about the m_i 's which can be expressed, by suitable choice of the parameters a_i and b_i , as the product of independent hyperbinomial functions.

In each stratum the cost per observation is a known constant. Using squared error loss function, the prior Bayes risk is found for the linear function of interest,

$$\theta = \sum_{i=1}^k \frac{\lambda_i m_i}{N_i},$$

and the optimum allocation of sample sizes is found, the one for which the prior Bayes risk is minimum when the total sampling budget is fixed.

CHAPTER 4

A LIST OF ARTICLES FOR ADDITIONAL READING

4.1 Biometrika

Bartholomew and Bassett (1966), "A comparison of some Bayesian and frequentist inferences II", *Biometrika*, 53, 262-264.

Bohrer (1966), "On Bayes sequential design with two random variables", *Biometrika*, 53, 469-475.

Tiao and Box (1967), "Bayesian analysis of a 3-component hierarchical design model", *Biometrika*, 54, 109-125.

Brown (1967), "The two means problem - a secondary Bayes approach", *Biometrika*, 54, 85-91.

Rutherford (1967), "The empirical Bayes approach: estimating posterior quantiles", *Biometrika*, 54, 672-675.

Bartholomew (1965), "A comparison of some Bayesian and frequentist inferences", *Biometrika*, 52, 19-35.

Lemon and Krutchkoff (1969), "An empirical Bayes smoothing technique", *Biometrika*, 56, 361-365.

Wetherill (1961), "Bayesian sequential analysis", *Biometrika*, 48, 281-292.

Tiao and Tan (1965), "Bayesian analysis of random effect model in the analysis of variance I: Posterior distribution of variance components", *Biometrika*, 52, 37-53.

Tiao and Tan (1966), "Bayesian analysis of random effect models in the analysis of variance II: Effect of autocorrelated errors", *Biometrika*, 53, 477-495.

Box and Tiao (1962), "A further look at robustness via Bayes theorem", *Biometrika*, 49, 419-432.

Tiao and Zellner (1964), "Bayes theorem and the use of prior knowledge in regression analysis", *Biometrika*, 51, 219-230.

Tiao and Draper (1968), "Bayesian analysis of linear models with two random components with special reference to the balanced incomplete block design", *Biometrika*, 55, 101-117.

Clemmer and Krutchkoff (1968), "The use of empirical Bayes estimators in a linear regression model", *Biometrika*, 55, 525-534.

Tiao and Fienberg (1969), "Bayesian estimation of latent roots and vectors with special reference to the bivariate normal distribution", *Biometrika*, 56, 97-108.

Rutherford and Krutchkoff (1969), "Asymptotic optimality of empirical Bayes estimators", *Biometrika*, 56, 220-223.

Martz and Krutchkoff (1969), "Empirical Bayes estimators in a multiple linear regression model", *Biometrika*, 56, 367-374.

Stone (1969), "The role of experimental randomization in Bayesian statistics - finite sampling and two Bayesians", *Biometrika*, 56, 681-683.

Gabbert and Krutchkoff (1970), "Supplementary sample nonparametric empirical Bayes in a quality control situation", *Biometrika*, 57, 67-69.

Freeman (1970), "Optimal Bayesian sequential estimation of the median effective dose", *Biometrika*, 57, 79-89.

George (1971), "Evaluation of empirical Bayes estimators for small numbers of past samples", *Biometrika*, 58, 244-251.

Preston (1971), "An Empirical Bayes problem with a Markovian parameter", *Biometrika*, 58, 535-543.

Altham (1971), "Exact Bayesian analysis of an intraclass 2x2 table", *Biometrika*, 58, 679-680.

Efron and Morris (1972), "Empirical Bayes on vector observations: An extension of Stein's method", *Biometrika*, 59, 335-347.

Copas (1972), "Empirical Bayes methods and the repeated use of a standard", *Biometrika*, 59, 349-360.

Stone and Dawid (1972), "Un-Bayesian implications of improper Bayes inference in routine statistical problems", *Biometrika*, 59, 369-376.

Draper and Guttman (1968), "Some Bayesian stratified two phase sampling results", *Biometrika*, 55, 131-139.

Maritz (1967), "Empirical Bayes estimation for the Poisson distribution", *Biometrika*, 54, 367-374.

Smith (1973), "Bayes estimators in one way and two way models", *Biometrika*, 60, 319-328.

Davidson and Solomon (1973), "A Bayesian approach to paired comparison experimentation", *Biometrika*, 60, 477-488.

O'Hagan (1973), "Bayes estimation of a convex quadratic", *Biometrika*, 60, 565-572.

Watson (1974), "On Bayesian inference with incompletely specified prior distributions", *Biometrika*, 61, 193-200.

Harville (1974), "Bayesian inference for variance components using only error contrasts", *Biometrika*, 61, 383-394.

Brooks (1974), "Bayesian analysis of the two sample problem under the Lehmann alternations", *Biometrika*, 61, 501-520.

Leonard (1974), "A modification to the Bayes estimate for the mean of a normal distribution", *Biometrika*, 61, 627-628.

Goldstein (1974), "Approximate Bayesian inference with incompletely specified prior distributions", *Biometrika*, 61, 629-632.

Jones (1974), "A note on the Bayesian sequential estimation of a binomial parameter", *Biometrika*, 61, 642-650.

4.2 Technometrics

Dunsmore (1974), "The Bayesian predictive distribution in life testing models", *Technometrics*, 16, 455-463.

Posten (1963), "Robustness of uniform Bayesian encoding", *Technometrics*, 5, 121-125

Springer and Thompson (1968), "Bayesian confidence limits for reliability of redundant systems when tests are terminated at first failure", *Technometrics*, 10, 29-36.

Steck and Zimmer (1968), "The relationship between Neyman and Bayes confidence intervals for the hypergeometric parameter", *Technometrics*, 10, 199-203.

Draper and Guttman (1971), "Bayesian estimation of the binomial parameter", *Technometrics*, 13, 667-673.

Barnett (1972), "A Bayes sequential life test", *Technometrics*, 14, 453-468.

Pierce (1973), "Fiducial, frequency and Bayesian inference on reliability for the two parameter negative exponential distribution", *Technometrics*, 15, 249-253.

4.3 Journal of American Statistical Association

Novick and Grizzle (1965), "A Bayesian approach to the analysis of data from clinical trials," *Am. Stat. Ass. J.*, 60, 81-96.

Novick and Hall (1965), "A Bayesian indifference procedure", *Am. Stat. Ass. J.*, 60, 1104-1117.

Karson and Wroblewski (1970), "A Bayesian analysis of a binomial model with a partially informative category - Proceedings of the business and economics statistics section," Am. Stat. Ass. J., 65, 532-534.

Zellner and Tiao (1964), "Bayesian analysis of the regression model with autocorrelated models", Am. Stat. Ass. J., 59, 763-778.

Cornfield (1966), "A Bayesian test of some classical hypothesis - with applications to sequential clinical trials", Am. Stat. Ass. J., 61, 577-584.

Hoadley (1970), "A Bayesian look at inverse linear regression", Am. Stat. Ass. J., 65, 356-368.

Zellner (1971), "Bayesian and non-Bayesian analysis of Log-normal distribution and Log-normal regression", Am. Stat. Ass. J., 66, 327-328.

Efron and Morris (1972), "Limiting risk of Bayes and empirical Bayes estimators", Am. Stat. Ass. J., 67, 130-138.

Lockner and Basu (1972), "Bayesian analysis of two sample problem with incomplete data", Am. Stat. Ass. J., 67, 432-450.

Efron and Morris (1973), "Steins' estimation rule and its competitors - an empirical Bayes approach", Am. Stat. Ass. J., 68, 117-125.

Aldrich and Morton (1973), "Bayesian modification of the failure rate function", Am. Stat. Ass. J., 68, 483-484.

Shapiro (1973), "On computing Bayes risk", Am. Stat. Ass. J., 68, 485-492.

Bassin (1973), "A Bayesian optimal overhaul interval model for Weibull restoration process", Am. Stat. Ass. J., 68, 575-592.

Swami and Mehta (1973), "Bayesian analysis of error components regression models", Am. Stat. Ass. J., 68, 648-653.

Bogdanoff and Pierce (1973), "Bayes-fiducial inference for the Weibull distribution", Am. Stat. Ass. J., 68, 659-663.

Kaufman and King (1973), "A Bayesian analysis of non-response in dichotomous processes", Am. Stat. Ass. J., 68, 670-675.

Swami and Mehta (1973), "Bayesian analysis of a bivariate normal distribution with incomplete observations", Am. Stat. Ass. J., 68, 922-928.

Duncan (1974), "An empirical Bayes approach to scoring multiple choice tests in the misinformation model", Am. Stat. Ass. J., 69, 50-57.

Press and Yang (1974), "A Bayesian approach to second guessing 'undecided' respondents", *Am. Stat. Ass. J.*, 69, 58-68.

Ferreira (1975), "A Bayesian analysis of a switching regression model", *Am. Stat. Ass. J.*, 70, 370-384.

4.4 Annals of Mathematical Statistics

Guthrie and Johns (1959), "Bayes acceptance sampling procedures for large lots", *Ann. Math. Stat.*, 30, 896-925.

Geissor (1965), "Bayesian estimation in multivariate analysis", *Ann. Math. Stat.*, 60, 150-159.

Hall and Novick (1963), "A note on classical and Bayesian prediction intervals for location, scale and regression models", *Ann. Math. Stat.*, 34, 1619-1620.

Box and Tiao (1965), "Multiparameter problems from a Bayesian point of view", *Ann. Math. Stat.*, 36, 1468-1482.

Bloch and Watson (1967), "A Bayesian study of the multinomial distributions", *Ann. Math. Stat.*, 38, 1423-1435.

Duncan (1961), "Bayes rule for a common multiple comparisons problem and related student-t problems", *Ann. Math. Stat.*, 32, 1013-1033.

Strawderman (1971), "Proper Bayes minimax estimators of the multivariate normal mean", *Ann. Math. Stat.*, 42, 385-388.

Dickey (1968), "3 multidimensional integral identities with Bayesian applications", *Ann. Math. Stat.*, 39, 1615-1620.

Pierce and Bogdanoff (1971), "A note on Bayes- fiducial intervals for problems of location and scale", *Ann. Math. Stat.*, 42, 833-837.

Robbins (1964), "The empirical Bayes approach to statistical decision problems", *Ann. Math. Stat.*, 35, 1-20.

Johns (1957), "Non-parametric Bayes procedures", *Ann. Math. Stat.*, 28, 649-669.

Amster (1963), "A modified Bayes stopping rule", *Ann. Math. Stat.*, 34, 1404-1413.

Bickel and Yahav (1968), "Asymptotically optimal Bayes and minimax procedures in sequential estimation", *Ann. Math. Stat.*, 39, 442-456.

Lindley (1964), "Bayesian analysis of contingency tables", *Ann. Math. Stat.*, 35, 1622-1643.

4.5 Journal of Royal Statistical Society

Tiao and Zellner (1964), "On the Bayesian estimation of multivariate regression", J. Royal Stat. Soc., Ser. B, 26, 277-285.

Pratt (1965), "Bayesian interpretation of standard inference statements", J. Royal Stat. Soc., Ser. B, 27, 169-203.

Dunsmore (1968), "A Bayesian approach to calibration", J. Royal Stat. Soc., Ser. B, 30, 396-405.

Hartigan (1969), "Linear Bayes methods", J. Royal Stat. Soc., Ser. B, 31, 446-454.

Copas (1969), "Compound decisions and empirical Bayes", J. Royal Stat. Soc., Ser. B, 31, 397-425.

Good (1967), "A Bayesian significance test for multinomial distributions", J. Royal Stat. Soc., Ser. B, 29, 399-418.

Draper and Guttman (1969), "Subjective Bayesian models in sampling finite problems", J. Royal Stat. Soc., Ser. B, 31, 195-233.

Thatcher (1964), "Relations between Bayesian and confidence limits for predictions", J. Royal Stat. Soc., Ser. B, 26, 176-210.

Aitchison (1964), "Bayesian tolerance regions", J. Royal Stat. Soc., Ser. B, 26, 161-175.

Hald (1967), "Asymptotic properties of Bayesian single sampling plans", J. Royal Stat. Soc., Ser. B, 29, 162-173.

Altham (1969), "Exact Bayesian analysis of a 2x2 contingency table and Fishers 'exact' significance test", J. Royal Stat. Soc., Ser. B, 31, 261-269.

Lindley and Smith (1972), "Bayes estimates for the linear model", J. Royal Stat. Soc., Ser. B, 34, 1-41.

Dawid, Stone and Zidek (1973), "Marginalization paradoxes in Bayesian and structural inference", J. Royal Stat. Soc., Ser. B, 35, 189-233.

4.6 Miscellanea

Novick (1964), "On Bayesian logical probability", Research Bulletin, 64-22, Princeton-N.J., Educational Testing Service.

Novick and Hall (1963), "A Bayesian indifference procedure", Institute of Statistics, Mimeograph Series #345, University of North Carolina.

Hald (1964), "Bayesian single sampling attribute plans for discrete prior distributions", Inst. of Math. Stat., University of Copenhagen.

Raiffa and Schlaifer (1961), Applied Decision Theory, Harvard University.

Lindley (1965), "Introduction to probability and statistics from a Bayesian viewpoint", Part II, Inference, Cambridge University Press.

Good (1965), "The estimation of probabilities - An essay on modern Bayesian methods", Cambridge, M.I.T. Press.

Martin (1967), "Bayesian Decision problems and Markov chains", New York, John Wiley.

Elfring (1966), "Robustness of Bayes decisions against the choice of prior", Technical report #122, Stanford University, Dept. of Statistics.

Novick and Thayer (1969), "Comparison of Bayesian estimate of the true score", Educational Testing Service, Princeton, N.J., RB-69-74.

Soland (1969), "Bayesian analysis of the Weibull process with unknown shape and scale parameters", IEEE on reliability, R-18, No. 4, 181-184.

Bonis (1966), "Bayesian reliability demonstration plans", Proceedings of 5th aerospace reliability and maintainability conference", 861-870.

Grosh (1969), "Bayes one and two stage stratified sampling schemes for finite populations with beta-binomial prior distribution", Ph.D. dissertation, Kansas State University.

Anscombe (1963), "Bayesian inference concerning many parameters with reference to supersaturated designs", I.S.I. Bulletin, 34th Session, Ottawa.

BIBLIOGRAPHY

- [1] Ando, A. and Kaufman, G. M. (1965), "Bayesian analysis of the independent multinormal process" - neither mean nor precision known", Am. Stat. Ass. J., 60, 347-358. [3.7]
- [2] Antelman, G. R. (1965), "Insensitivity to non-optimal design in Bayesian decision theory", Am. Stat. Ass. J., 60, 584-601. [3.8]
- [3] Behboodian, J. (1972), "Bayesian estimation for the proportions in a mixture of distributions", Sankhya, Ser. B., 34, 15-20.
- [4] Bennett, G. K. and Martz, H. F. (1972), "A continuous empirical Bayes smoothing technique", Biometrika, 59, 361-368. [3.34]
- [5] Bernard, G. A. (1954), "Sampling inspection and statistical decisions", J. R. Stat. Soc., B, 16, 151-174.
- [6] Bhattacharya, S. K. (1967), "Bayesian approach to life testing and reliability estimation", Am. Stat. Ass. J., 62, 48-62. [3.11]
- [7] Boes, D. C. (1966), "On the estimation of mixing distributions", Ann. Math. Stat., 37, 177-188.
- [8] Box, G. E. P. and Tiao, G. C. (1968), "Bayesian estimation of means for the random effect model", Am. Stat. Ass. J., 63, 174-181. [3.13]
- [9] Bracken, J. (1966), "Percentage points of the beta distribution for use in Bayesian analysis of Bernoulli processes", Technometrics, 8, 687-694. [3.45]
- [10] Bratcher, T. L.; Schcany, W. R. and Hunt, H. H. (1971), "Bayesian prediction and population size assumptions", Technometrics, 13, 678-681.
- [11] Cacoullos, T. (1965), "Comparing Mahalanobis distances II: Bayes procedures when the mean vectors are unknown", Sankhya, Ser. A, 27-28, 23-32. [3.2].
- [12] Chiu, W. K. (1974), "A new prior distribution for attributes sampling", Technometrics, 16, 93-102.
- [13] Cochran, W. G. (1963), "Sampling techniques", New York, John Wiley and Sons, Inc.
- [14] Cox, D. R. (1960), "Serial sampling acceptance schemes derived from Bayes's theorem", Technometrics, 2, 353-360.

- [15] Dayananda, R. A. and Evans, I. G. (1973), "Bayesian acceptance-sampling schemes for two sided tests of the mean of a normal distribution of known variance", *Am. Stat. Ass. J.*, 68, 131-136. [3.42]
- [16] Dodge, H. F. and Romig, H. G. (1944), "Sampling inspection tables", 106, John Wiley, N.Y.
- [17] Draper, N. R. and Guttman, I. (1966), "Unequal group variances in the fixed-effects one-way analysis of variance: A Bayesian sidelight", *Biometrika*, 53, 27-35. [3.23]
- [18] Draper, N. R. and Guttman, I. (1968), "Some Bayesian stratified two-phase sampling results", *Biometrika*, 55, 131-139. [3.30]
- [19] Duncan, D. B. (1965), "A Bayesian approach to multiple comparisons", *Technometrics*, 7, 171-222. [3.44]
- [20] El-Sayyad, G. M. and Freeman, P. R. (1973), "Bayesian sequential estimation of a Poisson process rate", *Biometrika*, 60, 289-296. [3.36]
- [21] Ericson, W. A. (1969), "Subjective Bayesian models in sampling finite populations", *J. Royal Stat., Soc., Ser. B*, 31, 1975-224.
- [22] Feller, W. (1966), "An introduction to probability and its applications", John Wiley and Sons, Inc., N.Y., 1964.
- [23] Geisser, S. (1965), "A Bayes approach for combining correlated estimates", *Am. Stat. Ass. J.*, 60, 602-607. [3.9]
- [24] Geisser, S. (1970), "Bayesian analysis of growth curve", *Sankhya, Ser. A*, 32, 53-64. [3.5]
- [25] Girshick, M. A. and Rubin, H. (1952), "A Bayes approach to a quality control model", *Ann. Math. Stat.*, 23, 114-125. [3.20]
- [26] Good, I. J. and Crook, J. F. (1974), "The Bayes/non-Bayes compromise and the multinomial distribution", *Am. Stat. Ass. J.*, 69, 711-720. [3.43]
- [27] Gould, M. W. (1959), "Combinatorial Identities", NSS 14, Contract DA 18-110-SC-294, p. 19.
- [28] Griffin, B. S. and Krutchkoff, R. G. (1971), "An empirical Bayes estimator for P successes in the binomial distributions", *Sankhya, Ser. B*, 33, 217-224.
- [29] Griffin, B. S. and Krutchkoff, R. G. (1971), "Optimal linear estimators: an empirical Bayes version with application to the binomial distribution", *Biometrika*, 58, 195-201, [3.33]

- [30] Grosh, D. L. (1972), "A Bayes sampling allocation scheme for stratified finite populations with hyperbinomial prior distributions", *Technometrics*, 14, 599-612. [3.47]
- [31] Gunel, E. and Dickey, J. (1974), "Bayes factors for independence in contingency tables", *Biometrika*, 61, 545-557. [3.39]
- [32] Hald, A. (1960), "The compound hypergeometric distribution and a system of single sampling inspection plans based on prior distributions and costs", *Technometrics*, 2, 275-340.
- [33] Hald, A. (1965), "Bayesian single sampling attribute plans for discrete prior distributions", *Mat. Fys. Skr. Dan. Vid. Selsk.*, 3.
- [34] Hald, A. (1968), "Bayesian single sampling attribute plans for continuous prior distributions", *Technometrics*, 10, 667-683.
- [35] Hartley, H. O. and Rao, J. N. K. (1968), "A new estimation theory for sampling surveys", *Biometrika*, 55, 547-558.
- [36] Hill, B. M. (1968), "Posterior distribution of percentiles: Bayes' theorem for sampling from a population", *Am. Stat. Ass. J.*, 63, 677-691. [3.14]
- [37] Hoadley, B. (1969), "The compound multinomial distribution and Bayesian analysis of categorical data from finite populations", *Am. Stat. Ass. J.*, 64, 216-229. [3.17]
- [38] Hoadley, B. (1970), "A Bayesian look at inverse linear regression", *Am. Stat. Ass. J.*, 65, 356-369. [3.19]
- [39] Kappenman, R. F., Geisser, S. and Antle, C. E. (1970), "Bayesian and fiducial solutions for the Fieller-Creasy problem", *Sanshya, Ser. B*, 31-32, 331-340. [3.4]
- [40] Krutchkoff, R. G. (1967), "A supplementary sample non-parametric empirical Bayes approach to some statistical decision problems", *Biometrika*, 54, 451-458. [3.28]
- [41] Lee, T. C., Judge, G. G. and Zellner, A. (1968), "Maximum likelihood and Bayesian estimation of transition probabilities", *Am. Stat. Ass. J.*, 63, 1162-1179. [3.15]
- [42] Lemon, G. H. and Krutchkoff, R. G. (1969), "An empirical Bayes smoothing technique", *Biometrika*, 56, 361-365.
- [43] Leonard, T. (1972), "Bayesian methods for binomial data", *Biometrika*, 59, 581-589. [3.35]
- [44] Leonard, T. (1973), "A Bayesian method for histograms", *Biometrika*, 60, 297-308. [3.37]

- [45] Lindley, D. V. (1964), "The Bayesian analysis of contingency tables", *Ann. Math. Stat.*, 35, 1622-1643. [3.21]
- [46] Maritz, J. S. (1966), "Smooth empirical Bayes estimation for one-parameter discrete distributions", *Biometrika*, 53, 417-429. [3.24]
- [47] Maritz, J. S. (1967), "Smooth empirical Bayes estimation for continuous distributions", *Biometrika*, 54, 435-450. [3.27]
- [48] Maritz, J. S. (1968), "On the smooth empirical Bayes approach to testing of hypothesis and the compound decision problem", *Biometrika*, 55, 83-100. [3.29]
- [49] Maritz, J. S. (1969), "Empirical Bayes estimation for the Poisson distribution", *Biometrika*, 56, 349-359, [3.32]
- [50] Martz, H. F. and Lian, M. G. (1974), "Empirical Bayes estimation of the binomial parameter", *Biometrika*, 61, 517-523. [3.38]
- [51] Novick, M. R. and Grizzle, J. E. (1965), "A Bayesian approach to the analysis of data from clinical trials", *Am. Stat. Ass. J.*, 60, 81-96. [3.6]
- [52] Novick, M. R. and Hall, W. J. (1965), "A Bayesian indifference procedure", *Am. Stat. Ass. J.*, 60, 1104-1117. [3.10]
- [53] Peach, P. and Littauer, S. B. (1946), "A note on sampling inspection", *Ann. Math. Stat.*, 17, 81-84.
- [54] Raiffa, H. and Schlaifer, R. (1961), "Applied Statistical decision theory", M.I.T. Press, p. 240.
- [55] Rao, J. N. K. and Ghangurde, P. D. (1972), "Bayesian optimization in sampling finite populations", *Am. Stat. Ass. J.*, 67, 439-443. [3.40]
- [56] Romberg, H. F. (1972), "Continuous sequential testing of a Poisson process to minimize the Bayes risk", *Am. Stat. Ass. J.*, 67, 921-926. [3.41]
- [57] Rutherford, J. R. and Krutchkoff, R. G. (1967), "The empirical Bayes approach: estimating the prior distribution", *Biometrika*, 54, 326-328. [3.26]
- [58] Rutherford, J. R. and Krutchkoff, R. G. (1969), "Some empirical Bayes techniques in point estimation", *Biometrika*, 56, 133-137, [3.31]
- [59] Rubin, H. and Sethuraman, J. (1965), "Bayes risk efficiency", *Sankhya, Ser. A.*, 27-28, 347-356. [3.1]

- [60] Samuel, E. (1963), "An empirical Bayes approach to the testing of certain parametric hypothesis", *Ann. Math. Stat.*, 34, 1370-1385. [3.22]
- [61] Scott, A. and Smith, T. M. F. (1971), "Bayes estimates for subclasses in stratified sampling", *Am. Stat. Ass. J.*, 66, 834-836.
- [62] Springer, M. D. and Thompson, W. E. (1966), "Bayesian confidence limits for the product of N binomial parameters", *Biometrika*, 53, 611-613. [3.25]
- [63] Steck, G. P. and Zimmer, W. J. (1968), "The relationship between Neyman and Bayes confidence intervals for the hypergeometric parameter", *Technometrics*, 10, 199-203.
- [64] Van Ryzin, J. (1966), "Bayes risk consistency of classification procedures using density estimation", *Sankhya, Ser. A*, 27-28, 261-270. [3.3]
- [65] Waller, R. A. and Duncan, D. B. (1969), "A Bayes rule for the symmetric multiple comparisons problem", *Am. Stat. Ass. J.*, 64, 1484-1503. [3.18]
- [66] Wetherill, G. B. (1960), "Some remarks on the Bayesian solution of the single sample inspection scheme", *Technometrics*, 341-352.
- [67] Winkler, R. L. (1967), "The assessment of prior distributions in Bayesian analysis", *Am. Stat. Ass. J.*, 62, 776-800. [3.12]
- [68] Zellner, A. and Tiao, G. C. (1964), "Bayesian analysis of the regression model with autocorrelated errors", *Am. Stat. Ass. J.*, 59, 763-778. [3.16]

BIBLIOGRAPHY (Additional Reading)

- [1] Aitchison (1964), "Bayesian tolerance regions", J. Royal Stat. Soc., Ser. B, 26, 161-175.
- [2] Aldrich and Morton (1973), "Bayesian modification of the failure rate function", Am. Stat. Ass. J., 68, 483-484.
- [3] Altham (1971), "Exact Bayesian analysis of an intraclass 2x2 table", Biometrika, 58, 679-680.
- [4] Altham (1969), "Exact Bayesian analysis of a 2x2 contingency table and Fishers 'exact' significance test", J. Royal Stat. Soc., Ser. B, 31, 261-269.
- [5] Amster (1963), "A modified Bayes stopping rule", Ann. Math. Stat., 34, 1404-1413.
- [6] Anscombe (1963), "Bayesian inference concerning many parameters with reference to supersaturated designs", I.S.I. Bulletin, 34th Session, Ottawa.
- [7] Barnett (1972), "A Bayes sequential life test", Technometrics, 14, 453-468.
- [8] Bartholomew (1965), "A comparison of some Bayesian and frequentist inferences", Biometrika, 52, 19-35.
- [9] Bartholomew and Bassett (1966), "A comparison of some Bayesian and frequentist inferences", Biometrika, 53, 262-264.
- [10] Bassin (1973), "A Bayesian optimal overhaul interval model for Weibull restoration process", Am. Stat. Ass. J., 68, 575-592.
- [11] Bickel and Yahav (1968), "Asymptotically optimal Bayes and minimax procedures in sequential estimation", Ann. Math. Stat., 39, 442-456.
- [12] Bloch and Watson (1967), "A Bayesian study of the multinomial distributions", Ann. Math. Stat., 38, 1423-1435.
- [13] Bogdanoff and Pierce (1973), "Bayes-fiducial inference for the Weibull distribution", Am. Stat. Ass. J., 68, 659-663.
- [14] Bohrer (1966), "On Bayes sequential design with two random variables", Biometrika, 53, 469-475.
- [15] Bonis (1966), "Bayesian reliability demonstration plans", Proceedings of 5th aerospace reliability and maintainability conference, 861-870.

- [16] Box and Tiao (1962), "A further look at robustness via Bayes theorem", *Biometrika*, 49, 419-432.
- [17] Box and Tiao (1965), "Multiparameter problems from a Bayesian point of view", *Ann. Math. Stat.*, 36, 1468-1482.
- [18] Brooks (1974), "Bayesian analysis of the two sample problem under the Lehmann alternations", *Biometrika*, 61, 501-520.
- [19] Brown (1967), "The two means problem - a secondary Bayes approach", *Biometrika*, 54, 85-91.
- [20] Clemmer and Krutchkoff (1968), "The use of empirical Bayes estimators in a linear regression model", *Biometrika*, 55, 525-534.
- [21] Copas (1972), "Empirical Bayes methods and the repeated use of a standard", *Biometrika*, 59, 349-360.
- [22] Copas (1969), "Compound decisions and empirical Bayes", *J. Royal Stat. Soc., Ser. B*, 31, 397-425.
- [23] Cornfield (1966), "A Bayesian test of some classical hypotheses - with application to sequential clinical trials", *Am. Stat. Ass. J.*, 61, 577-584.
- [24] Davidson and Solomon (1973), "A Bayesian approach to paired comparison experimentation", *Biometrika*, 60, 477-488.
- [25] Dawid, Stone and Zidek (1973), "Marginalization paradoxes in Bayesian and structural inference", *J. Royal Stat. Soc., Ser. B*, 35, 189-233.
- [26] Dickey (1968), "3 multidimensional integral identities with Bayesian applications", *Ann. Math. Stat.*, 39, 1615-1620.
- [27] Duncan (1961), "Bayes rule for a common multiple comparisons problem and related student -t problems", *Ann. Math. Stat.*, 32, 1013-1033.
- [28] Duncan (1974), "An empirical Bayes approach to scoring multiple choice tests in the misinformation model", *Am. Stat. Ass. J.*, 69, 50-57.
- [29] Dunsmore (1974), "The Bayesian predictive distribution in life testing models", *Technometrics*, 16, 455-463.
- [30] Dunsmore (1968), "A Bayesian approach to calibration", *J. Royal Stat. Soc., Ser. B*, 30, 396-405.
- [31] Draper and Guttman (1969), "Subjective Bayesian models in sampling finite problems", *J. Royal Stat. Soc., Ser. B*, 31, 195-233.

- [32] Draper and Guttman (1971), "Bayesian estimation of the binomial parameter", *Technometrics*, 13, 667-673.
- [33] Draper and Guttman (1968), "Some Bayesian stratified two phase sampling results", *Biometrika*, 55, 131-139.
- [34] Efron and Morris (1972), "Empirical Bayes on vector observations: An extension of Stein's method", *Biometrika*, 59, 335-347.
- [35] Efron and Morris (1972), "Limiting risk of Bayes and empirical Bayes estimators", *Am. Stat. Ass. J.*, 67, 130-138.
- [36] Efron and Morris (1973), "Stein's estimation rule and its competitors - an empirical Bayes approach", *Am. Stat. Ass. J.*, 68, 117-125.
- [37] Elfring (1966), "Robustness of Bayes decisions against the choice of prior", Technical report #122, Stanford University, Dept. of Statistics.
- [38] Ferreira (1975), "A Bayesian analysis of a switching regression model", *Am. Stat. Ass. J.*, 70, 370-384.
- [39] Freeman (1970), "Optimal Bayesian sequential estimation of the median effective dose", *Biometrika*, 57, 79-89.
- [40] Gabbert and Krutchkoff (1970), "Supplementary sample nonparametric empirical Bayes in a quality control situation", *Biometrika*, 57, 67-69.
- [41] Geissor (1965), "Bayesian estimation in multivariate analysis", *Ann. Math. Stat.*, 60, 150-159.
- [42] George (1971), "Evaluation of empirical Bayes estimators for small numbers of past samples", *Biometrika*, 58, 244-251.
- [43] Goldstein (1974), "Approximate Bayesian inference with incompletely specified prior distributions", *Biometrika*, 61, 629-632.
- [44] Good (1965), "The estimation of probabilities - An essay on modern Bayesian methods", Cambridge, M.I.T. Press.
- [45] Good (1967), "A Bayesian significance test for multinomial distributions", *J. Royal Stat. Soc., Ser. B*, 29, 399-418.
- [46] Grosh (1969), "Bayes one and two stage stratified sampling schemes for finite populations with beta-binomial prior distribution", Ph.D. dissertation, Kansas State University.

- [47] Guthrie and Johns (1959), "Bayes acceptance sampling procedures for large lots", *Ann. Math. Stat.*, 30, 896-925.
- [48] Hald (1964), "Bayesian single sampling attribute plans for discrete Prior distributions", *Inst. of Math. Stat., University of Copenhagen*.
- [49] Hald (1967), "Asymptotic properties of Bayesian single sampling plans", *J. Royal Stat. Soc., Ser. B*, 29, 162-173.
- [50] Hall and Novick (1963), "A note on classical and Bayesian prediction intervals for location, scale and regression models", *Ann. Math. Stat.*, 34, 1619-1620.
- [51] Hartigan (1969), "Linear Bayes methods", *J. Royal Stat. Soc., Ser. B*, 31, 446-454.
- [52] Harville (1974), "Bayesian inference for variance components using only error contrasts", *Biometrika*, 61, 383-394.
- [53] Hoadley (1970), "A Bayesian look at inverse linear regression", *Am. Stat. Ass. J.*, 65, 356-368.
- [54] Jones (1974), "A note on the Bayesian sequential estimation of a binomial parameter", *Biometrika*, 61, 642-650.
- [55] Johns (1957), "Non-parametric Bayes procedures", *Ann. Math. Stat.*, 28, 649-669.
- [56] Karson and Wroblewski (1970), "A Bayesian analysis of a binomial model with a partially informative category - Proceedings of the business and economics statistics section", *Am. Stat. Ass. J.*, 65, 532-534.
- [57] Kaufman and King (1973), "A Bayesian analysis of non-response in dichotomous processes", *Am. Stat. Ass. J.*, 68, 670-675.
- [58] Lemon and Krutchkoff (1969), "An empirical Bayes smoothing technique", *Biometrika*, 56, 361-365.
- [59] Leonard (1974), "A modification to the Bayes estimate for the mean of a normal distribution", *Biometrika*, 61, 627-628.
- [60] Lindley (1965), "Introduction to probability and statistics - from a Bayesian viewpoint", Part II, Inference, Cambridge University Press.
- [61] Lindley (1964), "Bayesian analysis of contingency tables", *Ann. Math. Stat.*, 35, 1622-1643.
- [62] Lindley and Smith (1972), "Bayes estimates for the linear model", *J. Royal Stat. Soc., Ser. B*, 34, 1-41.

- [63] Lockner and Basu (1972), "Bayesian analysis of two sample problem with incomplete data", *Am. Stat. Ass. J.*, 67, 432-450.
- [64] Maritz (1967), "Empirical Bayes estimation for the Poisson distribution", *Biometrika*, 54, 367-374.
- [65] Martin (1967), "Bayesian decision problems and Markov chains", New York, John Wiley.
- [66] Martz and Krutchkoff (1969), "Empirical Bayes estimators in a multiple linear regression model", *Biometrika*, 56, 367-374.
- [67] Novick (1964), "On Bayesian logical probability", *Research Bulletin*, 64-22, Princeton-N.J., Educational Testing Service.
- [68] Novick and Grizzle (1965), "A Bayesian approach to the analysis of data from clinical trials", *Am. Stat. Ass. J.*, 60, 81-96.
- [69] Novick and Hall (1965), "A Bayesian indifference procedure", *Am. Stat. Ass. J.*, 60, 1104-1117.
- [70] Novick and Hall (1963), "A Bayesian indifference procedure", Institute of Statistics, Mimeograph series #345, University of North Carolina.
- [71] Novick and Thayer (1969), "Comparison of Bayesian estimate of the true score", Educational Testing Service, Princeton, N.J., RB-69-74.
- [72] O'Hagan (1973), "Bayes estimation of a convex quadratic", *Biometrika*, 60, 565-572.
- [73] Pierce (1973), "Fiducial, frequency and Bayesian inference on reliability for the two parameter negative exponential distribution," *Technometrics*, 15, 249-253.
- [74] Pierce and Bogdanoff (1971), "A note on Bayes- fiducial intervals for problems of location and scale", *Ann. Math. Stat.*, 42, 833-837.
- [75] Postern (1963), "Robustness of uniform Bayesian encoding", *Technometrics*, 5, 121-125.
- [76] Pratt (1965), "Bayesian interpretation of standard inference statements", *J. Royal Stat. Soc., Ser. B*, 27, 169-203.
- [77] Press and Yang (1974), "A Bayesian approach to second guessing 'undecided' respondents", *Am. Stat. Ass. J.*, 69, 58-68.

- [78] Preston (1971), "An empirical Bayes problem with a Markovian parameter", *Biometrika*, 58, 535-543.
- [79] Raiff and Schlaifer (1961), *Applied Decision Theory*, Harvard University.
- [80] Robbins (1964), "The empirical Bayes approach to statistical decision problems", *Ann. Math. Stat.*, 35, 1-20.
- [81] Rutherford (1967), "The empirical Bayes approach: estimating posterior quantiles", *Biometrika*, 54, 672-675.
- [82] Rutherford and Krutchkoff (1969), "Asymptotic optimality of empirical Bayes estimators", *Biometrika*, 56, 220-223.
- [83] Shapiro (1973), "On computing Bayes risk", *Am. Stat. Ass. J.*, 68, 485-492.
- [84] Smith (1973), "Bayes estimators in one way and two way models", *Biometrika*, 60, 319-328.
- [85] Soland (1969), "Bayesian analysis of the Weibull process with unknown shape and scale parameters", *IEEE on reliability*, R-18, No. 4, 181-184.
- [86] Springer and Thompson (1968), "Bayesian confidence limits for reliability of redundant systems when tests are terminated at first failure", *Technometrics*, 10, 29-36.
- [87] Steck and Zimmer (1968), "The relationship between Neyman and Bayes confidence intervals for the hypergeometric parameter", *Technometrics*, 10, 199-203.
- [88] Stone (1969), "The role of experimental randomization in Bayesian statistics - finite sampling and two Bayesians", *Biometrika*, 56, 681-683.
- [89] Stone and Dawid (1972), "Un-Bayesian implications of improper Bayes inference in routine statistical problems", *Biometrika*, 59, 369-376.
- [90] Strawderman (1971), "Proper Bayes minimax estimators of the multivariate normal mean", *Ann. Math. Stat.*, 42, 385-388.
- [91] Swami and Mehta (1973), "Bayesian analysis of error components regression models", *Am. Stat. Ass. J.*, 68, 648-653.
- [92] Swami and Mehta (1973), "Bayesian analysis of a bivariate normal distribution with incomplete observations", *Am. Stat. Ass. J.*, 68, 922-928.
- [93] Thatcher (1964), "Relations between Bayesian and confidence limits for predictions", *J. Royal Stat. Soc., Ser. B*, 26, 176-210.

- [94] Tiao and Box (1967), "Bayesian analysis of a 3- component hierarchical design model", *Biometrika*, 54, 109-125.
- [95] Tiao and Draper (1968), "Bayesian analysis of linear models with two random components with special reference to the balanced incomplete block design", *Biometrika*, 55, 101-117.
- [96] Tiao and Fienberg (1969), "Bayesian estimation of latent roots and vectors with special reference to the bivariate normal distribution", *Biometrika*, 56, 97-108.
- [97] Tiao and Tan (1965), "Bayesian analysis of random effect models in the analysis of variance I: Posterior distribution of variance components", *Biometrika*, 52, 37-53.
- [98] Tiao and Tan (1966), "Bayesian analysis of random effect model in the analysis of variance II: Effect of autocorrelated errors", *Biometrika*, 53, 477-495.
- [99] Tiao & Zellner (1964), "Bayes theorem and the use of prior knowledge in regression analysis", *Biometrika*, 51, 219-230.
- [100] Tiao and Zellner (1964), "On the Bayesian estimation of multivariate regression", *J. Royal Stat. Soc., Ser. B*, 26, 277-285.
- [101] Watson (1974), "On Bayesian inference with incompletely specified prior distributions", *Biometrika*, 61, 193-200.
- [102] Wetherill (1961), "Bayesian sequential analysis", *Biometrika*, 48, 281-292.
- [103] Zellner (1971), "Bayesian and non-Bayesian analysis of Log-normal distribution and Log-normal regression", *Am. Stat., Ass. J.*, 66, 327-338.
- [104] Zellner and Tiao (1964), "Bayesian analysis of the regression model with autocorrelated models", *Am. Stat. Ass. J.*, 59, 763-778.

ACKNOWLEDGEMENT

The author of this report wishes to express his sincere thanks and appreciation to his major professor Dr. Doris Grosh for her helpful suggestions, encouragement and valuable criticism given to him during the preparation of this report.

BAYESIAN APPROACH TO QUALITY CONTROL

by

PRAKASH VAMAN JOSHI

B.E. (Electrical)

University of Poona, India (1974)

AN ABSTRACT OF A REPORT

submitted in partial fulfillment
of the requirements for the degree

MASTER OF SCIENCE

Department of Industrial Engineering

KANSAS STATE UNIVERSITY

Manhattan, Kansas

1976

ABSTRACT

The main purpose of this report is to see how Bayesian analysis is applicable in Quality Control. In all, this report contains 8 articles dealing with various aspects of Quality Control such as optimum sample size, acceptance number, costs for accepted or rejected lots, subclasses in stratified sampling, population size assumption, single sample inspection schemes etc. To quote a few results, it has been found out that optimum sample size is directly proportional to the square root of the lot size; it is also shown that for the prediction problem, the population may be assumed to be infinite or finite.

The report also gives a short summary of articles published by various authors in various journals that are concerned with Bayesian analysis. A list of articles for additional reading has also being given. In the end a bibliography is given.

This report covers the following journals over a period of 10 to 12 years. (1) Biometrika, (2) Technometrics, (3) Journal of Royal Statistical Society, (4) Journal of American Statistical Association, (5) Annals of Mathematical Statistics, (6) Journal of Quality Technology and (7) other miscellaneous journals which were cited.

Initially it was decided to cover a wide range of literature on Bayesian analysis, but due to the advanced and complicated nature of the subject, together with the lack of sufficient background, work has been severely restricted. Hence more complicated articles are either summarized or listed for additional reading.